

# CENTRAL LIMIT THEOREM FOR PARTIAL LINEAR EIGENVALUE STATISTICS OF WIGNER MATRICES

ZHIGANG BAO, GUANGMING PAN, AND WANG ZHOU

**ABSTRACT.** In this paper, we study the complex Wigner matrices  $M_n = \frac{1}{\sqrt{n}}W_n$  whose eigenvalues are typically in the interval  $[-2, 2]$ . Let  $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$  be the ordered eigenvalues of  $M_n$ . Under the assumption of four matching moments with the Gaussian Unitary Ensemble (GUE), for test function  $f$  4-times continuously differentiable on an open interval including  $[-2, 2]$ , we establish central limit theorems for two types of partial linear statistics of the eigenvalues. The first type is defined with a threshold  $u$  in the bulk of the Wigner semicircle law as  $\mathcal{A}_n[f; u] = \sum_{l=1}^n f(\lambda_l) \mathbf{1}_{\{\lambda_l \leq u\}}$ . And the second one is  $\mathcal{B}_n[f; k] = \sum_{l=1}^k f(\lambda_l)$  with positive integer  $k = k_n$  such that  $k/n \rightarrow y \in (0, 1)$  as  $n$  tends to infinity. Moreover, we derive a weak convergence result for a partial sum process constructed from  $\mathcal{B}_n[f; \lfloor nt \rfloor]$ .

## 1. INTRODUCTION

The complex Wigner Ensemble is defined as a family of  $n \times n$  random Hermitian matrices  $M_n$  of the form

$$M_n = \frac{1}{\sqrt{n}}W_n = \frac{1}{\sqrt{n}}\{w_{jk}\}_{j,k=1}^n,$$

in which  $w_{ll} \in \mathbb{R}, 1 \leq l \leq n$ ,  $w_{jk} = \bar{w}_{kj} \in \mathbb{C}, 1 \leq j < k \leq n$ , and  $\{w_{ll}, w_{jk}; 1 \leq l \leq n, 1 \leq j < k \leq n\}$  is a collection of independent variables such that

$$\mathbb{E}w_{ll} = \mathbb{E}w_{jk} = 0, \quad \mathbb{E}|w_{jk}|^2 = 1, \quad \mathbb{E}w_{ll}^2 = \sigma^2 < \infty.$$

Our basic additional assumption on the elements of  $W_n$  throughout the paper is the following condition.

*Condition  $\mathbf{C}_0$ :* We say that a complex Wigner matrix  $M_n$  obeys Condition  $\mathbf{C}_0$  if  $\{w_{ll}, \operatorname{Re}w_{jk}, \operatorname{Im}w_{jk}; 1 \leq l \leq n, 1 \leq j < k \leq n\}$  is a collection of independent variables whose distributions are all supported on at least three points, and we have the exponential decay condition on the elements in the sense that

$$\mathbb{P}(|w_{jk}| \geq t^C) \leq e^{-t}$$

---

*Date:* March 6, 2013.

2010 *Mathematics Subject Classification.* 15B52, 60F05, 60F17.

Z.G. Bao was partially supported by NSFC grant 11071213, NSFC grant 11101362, ZJNSF grant R6090034 and SRFDP grant 20100101110001; G.M. Pan was partially supported by the Ministry of Education, Singapore, under grant # ARC 14/11; W. Zhou was partially supported by the Ministry of Education, Singapore, under grant # ARC 14/11, and by a grant R-155-000-116-112 at the National University of Singapore.

holds for all  $t \geq C'$  with some positive constants  $C, C'$  (independent of  $j, k, n$ ).

A basic example of the complex Wigner matrix satisfying *Condition  $\mathbf{C}_0$*  is drawn from the Gaussian Unitary Ensemble (GUE) whose elements are Gaussian distributed, i.e.

$$w_{ll} \sim N(0, 1)_{\mathbb{R}}, \quad 1 \leq l \leq n, \quad w_{jk} \sim N(0, 1)_{\mathbb{C}}, \quad 1 \leq j < k \leq n.$$

Here  $N(0, 1)_{\mathbb{R}}$  (resp.  $N(0, 1)_{\mathbb{C}}$ ) represents the standard real (resp. complex) Gaussian distribution.

For Wigner matrix  $M_n$ , we denote its ordered eigenvalues as  $\lambda_1(M_n) \leq \lambda_2(M_n) \leq \dots \leq \lambda_n(M_n)$ . And the empirical spectral distribution (ESD) of  $M_n$  is defined by

$$F^{M_n}(x) =: \frac{1}{n} \sum_{l=1}^n \mathbf{1}_{\{\lambda_l(M_n) \leq x\}}.$$

When there is no confusion, we will briefly use  $\lambda_l$  and  $F_n(x)$  to represent  $\lambda_l(M_n)$  and  $F^{M_n}(x)$  respectively.

A fundamental result is the Wigner semicircle law, which describes the global limiting behavior of eigenvalues of the Wigner ensemble: for any  $\varphi \in C_b(\mathbb{R})$  (the set of bounded continuous functions in  $\mathbb{R}$ ), one has

$$\frac{1}{n} \sum_{l=1}^n \varphi(\lambda_l) \xrightarrow{\mathbb{P}} \int \varphi(x) \rho_{sc}(x) dx, \quad (1.1)$$

where

$$\rho_{sc}(x) = \frac{1}{2\pi^2} \sqrt{4 - x^2} \mathbf{1}_{\{|x| \leq 2\}}$$

is the density function of the Wigner semicircle law  $F_{sc}(x)$ . That is to say, the ESD  $F_n(x)$  converges weakly in probability to the semicircle law  $F_{sc}(x)$ . We remark here (1.1) holds under much weaker condition than  $\mathbf{C}_0$  assumed in this paper, see [2] for instance. Note that (1.1) can be viewed as a universal result corresponding to the classical law of large number (LLN) for sums of independent random variables. The quantity

$$\mathcal{L}_n[\varphi] = \sum_{l=1}^n \varphi(\lambda_l)$$

is usually referred to as the global linear eigenvalue statistic (**GLS**) of Wigner matrices with test function  $\varphi$ .

Once the LLN was obtained, a natural question in the probability theory is to study the fluctuation of  $\mathcal{L}_n[\varphi]$  subsequently. For any smooth enough test function  $\varphi$ , there are a vast of results obtained on the central limit theorem for  $\mathcal{L}_n[\varphi]$  under different assumptions, for instance, see [3], [4], [7], [22], [27], [28]. A remarkable work on this topic is due to Lytova and Pastur [22]. Particularly for GUE, Lytova

and Pastur showed that for any bounded test function  $\varphi$  with bounded derivative, one has

$$\mathcal{L}_n[\varphi] - \mathbb{E}\mathcal{L}_n[\varphi] \xrightarrow{d} N(0, V_{GUE}[\varphi]), \quad (1.2)$$

where

$$V_{GUE}[\varphi] = \frac{1}{4\pi^2} \int_{-2}^2 \int_{-2}^2 \left( \frac{\varphi(\lambda) - \varphi(\mu)}{\lambda - \mu} \right)^2 \frac{4 - \lambda\mu}{\sqrt{4 - \lambda^2} \sqrt{4 - \mu^2}} d\lambda d\mu. \quad (1.3)$$

(See Remark 2.1 of [22]). Moreover, an analogous result for more general Wigner matrices can be derived through the discussion in [22] for essentially  $C^5$  test functions. As the reader might notice, there is no normalizing constant in the convergence in (1.2). The reason behind this is mainly that the eigenvalues repel each other and so are more regularly distributed than independent random variables.

The main aim of this paper is to study the CLTs for two types of partial linear eigenvalue statistics (**PLES**) in the sense that only a part of eigenvalues will be involved in the statistics. The type 1 **PLES** with the test function  $f$  and the threshold  $u \in [-2 + \delta, 2 - \delta]$  is defined by

$$\mathcal{A}_n[f; u] = \sum_{l=1}^n f(\lambda_l) \mathbf{1}_{\{\lambda_l \leq u\}},$$

which is a summation of  $f(\lambda_l)$  only for  $\lambda_l \leq u$ . The type 2 **PLES** with the test function  $f$  and the integer  $k =: k_n$  is defined by

$$\mathcal{B}_n[f; k] = \sum_{l=1}^k f(\lambda_l),$$

with the constraint that  $k/n \rightarrow y \in (0, 1)$  as  $n$  tends to infinity. Note that the type 1 **PLES**  $\mathcal{A}_n[f; u]$  is just a **GLES** with the probably discontinuous test function  $f(x) \mathbf{1}_{\{x \leq u\}}$ . When  $f(u) = 0$ ,  $f(x) \mathbf{1}_{\{x \leq u\}}$  is continuous but may be non differentiable at  $u$ . We remark here though we define the two types of **PLES** as the sum of  $f(\lambda_l)$  for the smallest eigenvalues, it will cause no intrinsic difference on all discussions throughout the paper if we define the **PLES** with the largest eigenvalues instead of the smallest ones.

**PLES** for a variety of matrices (deterministic or random) play relevant roles in a lot of fields. For instance, when  $f(x) \equiv 1$ , the type 1 **PLES** is just the counting function of the eigenvalues up to  $u$ , which is a fundamental and well studied quantity in Random Matrix Theory (RMT). For the fluctuation of the counting function of Wigner matrices, we refer to [8], [20], [29], [10] for details of this topic. For the type 2 **PLES**, a canonical example is the sum of the  $k$  largest or smallest eigenvalues, which is important in both pure and applied aspects of matrix theory. Especially, the sum of the  $k$  largest eigenvalues is interesting in a lot of fields such as principal component analysis, compressed sensing and computational mathematics, see [1], [5], [13], [24] for instance. However, the type 2 **PLES** are always not easy to be studied since they are concerned with the ordered eigenvalues. By the generalized Rayleigh-Ritz theorem (see Corollary 4.3.18 of [21] for instance), one has for an  $n \times n$

Hermitian matrix  $A$ , there exists a variational representation as

$$\sum_{l=1}^k \lambda_l(A) = \min_{U^*U=I_k} \text{Tr} U^* A U \quad (1.4)$$

with any  $1 \leq k \leq n$ . Here  $\lambda_1(A) \leq \lambda_2(A) \leq \dots \leq \lambda_n(A)$  are ordered eigenvalues of  $A$ . However, such a variational characterization is not convenient for computation and analysis. Instead, one can work with a proxy of the quantity (1.4) by a type 1 partial sum with a threshold  $u$  “near”  $\lambda_k(A)$  as

$$\mathcal{A}_n[x; u] = \sum_{l=1}^n \lambda_l(A) \mathbf{1}_{\{\lambda_l(A) \leq u\}}.$$

Such an elementary approximate technic suggests us to study the two types of **PLES** together. As will be seen, with the aid of the so-called rigidity property of the eigenvalues proved by Erdős, Yau and Yin [18], such an approximate strategy does work well for the study of the fluctuation of a type 2 **PLES**. As mentioned above, an advantage of  $\mathcal{A}_n[f; u]$  is that it is indeed a **GLES**, though the test function is not necessarily continuous. Such a fact can help one avoid working with ordered eigenvalues.

For brevity, we set

$$f_u(x) =: (f(x) - f(u)) \mathbf{1}_{\{x \leq u\}}.$$

And for  $t \in [0, 1]$ , let  $\gamma_t$  be the number that

$$\int_{-2}^{\gamma_t} \rho_{sc}(x) dx = t.$$

Moreover, for some small positive number  $\delta$ , we set the interval

$$\mathcal{U} = (-2 - \delta, 2 + \delta)$$

throughout the paper. We use the notation  $C^k(\mathcal{U})$  to indicate the set of the real functions which are defined on the whole real line and  $k$ -times continuously differentiable on the interval  $\mathcal{U}$ . Our first result is the following theorem.

**Theorem 1.1.** *If  $M_n$  is drawn from  $GUE$ ,  $u \in [-2 + \delta, 2 - \delta]$  with some small but fixed  $\delta > 0$  and  $k =: k_n$  such that  $k/n \rightarrow y$  for some fixed constant  $y \in (0, 1)$  as  $n$  tends to infinity, one has the following CLTs.*

(i): *If  $f \in C^1(\mathcal{U})$  and  $f(u) \neq 0$ , one has*

$$\frac{\mathcal{A}_n[f; u] - m[f; u]}{\sqrt{\frac{f^2(u)}{2\pi^2} \log n}} \xrightarrow{d} N(0, 1)$$

with

$$m[f; u] = n \int_{-2}^u f(x) dF_{sc}(x).$$

(ii): *If  $f \in C^4(\mathcal{U})$  and  $f(u) = 0$ , one has*

$$\mathcal{A}_n[f; u] - m[f; u] \xrightarrow{d} N(0, V_{GUE}[f_u])$$

(iii): If  $f \in C^4(\mathcal{U})$ , for  $\mathcal{B}_n[f; k]$  we have

$$\mathcal{B}_n[f; k] - m[f; \gamma_{k/n}] \xrightarrow{d} N(0, V_{GUE}[f_{\gamma_y}]).$$

**Remark 1.2.** If  $f(x)$  is continuous in  $\mathbb{R}$  and its magnitude grows no faster than  $C_1 e^{c_1 x^2}$  as  $|x| \rightarrow \infty$  for some constants  $c_1 > 0$  and  $C_1 < \infty$ , we can also replace  $m[f; u]$  and  $m[f; \gamma_{k/n}]$  by  $\mathbb{E}\mathcal{A}_n[f; u]$  and  $\mathbb{E}\mathcal{B}_n[f; k]$  respectively in the above formulas. This is a consequence of the fact that the density of  $\mathbb{E}F_n(x)$  has a tail of  $O(e^{-cnx^2})$ . For details, see the forthcoming discussions in Section 2 and Section 3.

**Remark 1.3.** Observe that  $\mathcal{A}_n[f; u] = \mathcal{L}_n[f_u]$  if  $f(u) = 0$ . Thus (ii) implies that the differentiability condition imposed on the test function is not necessary for Lytova and Pastur's CLT, since  $f_u$  may be non-differentiable at  $u$ . At the same time, one can also learn from (i) that a discontinuous point of the test function will indeed cause significant change on the fluctuation of **GLES**. Note that there is still a large gap between differentiability and discontinuity. It will be interesting to investigate the relation between the smoothness of the test function and the limiting behavior of the corresponding **GLES**.

For ease of presentation, we use the notation

$$\xi^\circ =: \xi - \mathbb{E}\xi$$

for any random variable  $\xi$  in the sequel. Note that (iii) of Theorem 1.1 reveals the weak convergence of the random sequence  $\{\mathcal{B}_n[f; k]\}$ . Inspired by the classical partial sum process of i.i.d random variables, we take a step further to study the following partial sum process constructed from  $\mathcal{B}_n[f; k]$  with some small but fixed  $\delta > 0$  as

$$\mathcal{S}_n[f; t] = \mathcal{B}_n[f; \lfloor nt \rfloor] + (nt - \lfloor nt \rfloor)f(\lambda_{\lfloor nt \rfloor + 1}), t \in [\delta, 1 - \delta],$$

which is an element of  $C[\delta, 1 - \delta]$ . Here  $C[a, b]$  represents the space of continuous functions on the interval  $[a, b]$  equipped with uniform topology. Then for  $\mathcal{S}_n^\circ[f; t]$ , we have the following weak convergence theorem in  $C[\delta, 1 - \delta]$ .

**Theorem 1.4.** Suppose that  $M_n$  is drawn from  $GUE$ , and  $f \in C^3(\mathcal{U})$ . We also assume that there exist constants  $c_1 > 0$  and  $C_1 < \infty$  such that  $f(x)$  is continuous in  $\mathbb{R}$  and its magnitude grows no faster than  $C_1 e^{c_1 x^2}$  as  $|x| \rightarrow \infty$ . Then the sequence  $(\mathcal{S}_n^\circ[f; t]; t \in [\delta, 1 - \delta])$  is tight and converges weakly to a Gaussian process  $(\mathcal{S}[f; t]; t \in [\delta, 1 - \delta])$  with mean zero and covariance function given by

$$\begin{aligned} & \text{Cov}(\mathcal{S}[f; s], \mathcal{S}[f; t]) \\ &= \frac{1}{4\pi^2} \int_{-2}^2 \int_{-2}^2 \left( \frac{f_{\gamma_t}(\lambda) - f_{\gamma_t}(\mu)}{\lambda - \mu} \right) \left( \frac{f_{\gamma_s}(\lambda) - f_{\gamma_s}(\mu)}{\lambda - \mu} \right) \frac{4 - \lambda\mu}{\sqrt{4 - \lambda^2} \sqrt{4 - \mu^2}} d\lambda d\mu. \end{aligned}$$

**Remark 1.5.** Actually, one can extend the above result to the test function  $f \in C^3(\mathcal{U})$  without any additional condition imposed on its growth as  $|x| \rightarrow \infty$  if we consider the process  $\mathcal{S}[f; t] - \mathbb{E}\mathcal{S}[\bar{f}_\epsilon; t]$  instead. Here  $\bar{f}_\epsilon(x)$  is a smooth truncation of  $f(x)$  in the sense that  $\bar{f}_\epsilon(x) =: \chi_\epsilon(x)f(x)$ , where  $\chi_\epsilon(x)$  is a smooth cutoff to the region  $|x| \leq 2 + \epsilon$  that equals 1 for  $|x| \leq 2 + \epsilon/2$  with some small positive

number  $\epsilon < \delta$ . Such an extension can be achieved easily by using the large deviation estimate of extreme eigenvalues (See Lemma 7.4 for instance). We leave the detail to the reader.

**Remark 1.6.** It would also seem natural to consider the process constructed from the type 1 PLES  $\mathcal{A}_n[f; u]$  with the parameter  $u \in [-2 + \delta, 2 - \delta]$ , which can be viewed as an element in  $D[-2 + \delta, 2 - \delta]$  (the metric space of functions on  $[-2 + \delta, 2 - \delta]$  with discontinuities of the first kind, equipped with Skorokhold metric.) However, we assert that for general  $f$ ,  $\mathcal{A}_n^\circ[f; u]$  cannot converge weakly to any non-trivial process for any choice of normalization. For instance, when  $f(x) \equiv 1$ , such a fact has been mentioned in Bai and Silverstein [2] for sample covariance matrices. The case of Wigner matrices is just analogous.

The next two results are the extensions of Theorem 1.1 and Theorem 1.4 to more general complex Wigner matrices. In order to state our results, we use the terminology of Tao and Vu (see [32] for example) to say that  $M_n = \frac{1}{\sqrt{n}}(w_{jk})_{j,k=1}^n$  matches  $M'_n = \frac{1}{\sqrt{n}}(w'_{jk})_{j,k=1}^n$  to the  $\beta$ -th order off the diagonal and the  $\gamma$ -th order on the diagonal if

$$\mathbb{E}(w_{ll})^\alpha = \mathbb{E}(w'_{ll})^\alpha, \quad 0 \leq \alpha \leq \gamma, \quad 1 \leq l \leq n,$$

$$\mathbb{E}(\text{Re} w_{jk})^{\alpha_1} (\text{Im} w_{jk})^{\alpha_2} = \mathbb{E}(\text{Re} w'_{jk})^{\alpha_1} (\text{Im} w'_{jk})^{\alpha_2}, \quad 0 \leq \alpha_1 + \alpha_2 \leq \beta, \quad 1 \leq j < k \leq n,$$

where  $\alpha_1, \alpha_2$  and  $\alpha$  are non-negative integers. We state our results as follows.

**Theorem 1.7.** Suppose that  $M_n$  is a Wigner matrix satisfying Condition  $\mathbf{C}_0$  and matches to GUE to the fourth order off the diagonal and the second order on the diagonal. Then for the test function  $f \in C^4(\mathcal{U})$ , (i) – (iii) of Theorem 1.1 still hold for  $M_n$ .

**Theorem 1.8.** Suppose that  $M_n$  is a Wigner matrix satisfying Condition  $\mathbf{C}_0$  and matches to GUE to the fourth order off the diagonal and the second order on the diagonal. We assume  $f \in C^4(\mathcal{U})$ . Additionally, we assume there exist constants  $K < \infty$  and  $C_1 < \infty$  independent of  $n$ , such that  $f(x)$  is continuous in  $\mathbb{R}$  and its magnitude grows no faster than  $C_1|x|^K$  when  $|x| \rightarrow \infty$ . Then we also have that the process  $(\mathcal{S}_n^\circ[f; t]; t \in [\delta, 1 - \delta])$  is tight and converges weakly to  $(\mathcal{S}[f; t]; t \in [\delta, 1 - \delta])$ .

**Remark 1.9.** Similar to Remark 1.5, one can also extend the above result to  $f \in C^4(\mathcal{U})$  if one considers the process  $\mathcal{S}[f; t] - \mathbb{E}\mathcal{S}[\bar{f}_\epsilon; t]$  instead. Moreover, it is likely that one can extend the above result to  $t \in [\delta, 1]$  with further discussion on the edge of the spectrum. However, the current issue relies on some crucial estimates proved only for the bulk case, such as Lemma 5.5. So we do not pursue this direction here.

From now on, we will use the notation  $C, C_1, C'$  and  $L$  to denote some  $n$ -independent positive constants whose values may differ from line to line. And throughout the paper, we say an event  $E$  holds with high probability if

$$P(E) \geq 1 - n^{-c}$$

with some constant  $c$  and with overwhelming probability if

$$P(E) \geq 1 - n^{-C}$$

for any constant  $C > 0$ .

Our paper is organized as follows. In Section 2, we provide some basic tools and preliminaries of the whole paper. And Section 3 is devoted to the proof of the CLTs for two types of **PLES** for GUE, i.e. Theorem 1.1. Then in Section 4, we prove Theorem 1.4, whose proof is heavily based on the discussion in Section 3. In Section 5, we prove a comparison theorem for linear eigenvalue statistics, see Theorem 5.1. And as an application, we use our comparison theorem to extend Theorem 1.1 to general complex Wigner matrices, i.e. Theorem 1.7. Also with the aid of the comparison theorem, we prove Theorem 1.8 in Section 6. Some necessary known results are stated in the Appendix.

## 2. PRELIMINARIES

In this section, we will provide some basic notions and tools necessary for our proof in the sequel. Totally speaking, our strategy is to prove the results for GUE first and then extend them to general Wigner matrices by some comparison procedure. Such a strategy is quite fundamental in RMT. Thus the basic tools presented in this section consist of two separated parts. The first part is particularly for GUE, and the second part will be mainly contributed to our comparison procedure.

Using GUE as our starting point is mainly because its explicit formula of the joint probability density (j.p.d.) for the eigenvalues has a determinantal structure, which is friendly with analysis. By making use of the j.p.d., a vast of central issues in RMT can be solved explicitly for GUE. We refer to the books of Deift [12] and Mehta [23] for comprehensive surveys in this aspect.

If  $M_n = \frac{1}{\sqrt{n}}W_n$  is drawn from GUE, then the joint distribution of non ordered eigenvalues of  $W_n$  has the following j.p.d.

$$\rho_{n,n}(x_1, \dots, x_n) = \frac{1}{n!} \prod_{1 \leq j < k \leq n} |x_j - x_k|^2 e^{-\sum_{j=1}^n x_j^2/2}.$$

For the point process  $x_1, \dots, x_n$ , the  $k$ -point correlation function  $\rho_{k,n}$  has the well known determinantal structure

$$\begin{aligned} \rho_{k,n}(x_1, \dots, x_k) &= \frac{n!}{(n-k)!} \int_{\mathbb{R}^{n-k}} \rho_{n,n}(x_1, \dots, x_n) dx_{k+1} \dots dx_n \\ &= \det(K_n(x_l, x_j))_{l,j=1}^k. \end{aligned}$$

Here  $K_n(x, y)$  is the kernel function given by

$$K_n(x, y) = \sum_{l=0}^{n-1} H_l(x) H_l(y) e^{-\frac{1}{4}(x^2+y^2)} = \sum_{i=0}^{n-1} \psi_i(x) \psi_i(y),$$

where  $H_l(x)$  is the  $l$ -th orthonormalized Hermite polynomial w.r.t. the weight function  $e^{-x^2/2}$  and  $\psi_l(x)$  is the corresponding oscillator wave function. By the famous Christoffel-Darboux formula, one has for  $x \neq y$

$$K_n(x, y) = \sqrt{n} \frac{\psi_n(x) \psi_{n-1}(y) - \psi_{n-1}(x) \psi_n(y)}{x - y}$$

and for  $x = y$  by l'Hôpital's rule,

$$K_n(x, x) = n\psi_{n-1}^2(x) - \sqrt{n(n-1)}\psi_{n-2}(x)\psi_n(x). \quad (2.1)$$

Using the notation

$$\mathcal{K}_n(x, y) = \sqrt{n}K_n(\sqrt{n}x, \sqrt{n}y),$$

one has the following explicit formulas of expectation and variance of  $\mathcal{L}_n[\varphi]$  for GUE.

$$\mathbb{E}\mathcal{L}_n[\varphi] = \int_{\mathbb{R}} \varphi(x)\mathcal{K}_n(x, x)dx \quad (2.2)$$

and

$$\text{Var}\mathcal{L}_n[\varphi] = \frac{1}{2} \int_{\mathbb{R}} \int_{\mathbb{R}} (\varphi(x) - \varphi(y))^2 \mathcal{K}_n^2(x, y) dx dy. \quad (2.3)$$

See (4) of [29] and (2.27) of [25] for instance. Note that from (2.2),  $\mathcal{K}_n(x, x)/n$  is just the density function of  $\mathbb{E}F_n(x)$ . For our purpose, we state below some properties for the kernel function  $\mathcal{K}_n(x, y)$ . Firstly, by definition and the Cauchy-Schwarz inequality,

$$|\mathcal{K}_n(x, y)|^2 \leq \mathcal{K}_n(x, x)\mathcal{K}_n(y, y) \quad (2.4)$$

Moreover, by (2.1), one has

$$\mathcal{K}_n(x, x) = \sqrt{n}(n\psi_{n-1}^2(\sqrt{n}x) - \sqrt{n(n-1)}\psi_{n-2}(\sqrt{n}x)\psi_n(\sqrt{n}x)).$$

By adjusting the scale in the setting of [20], one can see that when  $|x| \geq 2 + \varepsilon$  for some small  $\varepsilon > 0$ ,

$$\psi_n(\sqrt{n}x) = \mathcal{O}(n^{-1/4}e^{-\frac{n}{4}F(x)}), \quad (2.5)$$

where

$$F(x) = \left| \int_x^2 \sqrt{|4 - y^2|} dy \right|.$$

See Section 4 of [20] for reference. Consequently,

$$\frac{\mathcal{K}_n(x, x)}{n} = \mathcal{O}(e^{-cnx^2}) \quad (2.6)$$

for  $|x| \geq 2 + \varepsilon$ . Such a fact has been mentioned in Remark 1.2. By (2.2)-(2.6), one can see that when  $\varphi(x)$  is continuous and  $|\varphi(x)|$  grows more slowly than  $C_1 e^{c_1 x^2}$  as  $|x| \rightarrow \infty$  for some  $c_1 > 0$  and  $C_1 < \infty$ , we have for sufficiently large  $n$

$$\mathbb{E}\mathcal{L}_n[\varphi] = \int_{\mathcal{U}} \varphi(x)\mathcal{K}_n(x, x)dx + \mathcal{O}(e^{-cn}) \quad (2.7)$$

and

$$\text{Var}\mathcal{L}_n[\varphi] = \frac{1}{2} \int_{\mathcal{U}} \int_{\mathcal{U}} (\varphi(x) - \varphi(y))^2 \mathcal{K}_n^2(x, y) dx dy + \mathcal{O}(e^{-cn}) \quad (2.8)$$

with some positive constant  $c$  depending only on  $\varphi$  and  $\delta$ . The above formulas will be frequently used in our proof for the GUE case in Sections 3 and 4.

However, for general Wigner matrices, the explicit formula for the joint distribution of the eigenvalues is obviously not available. A classical strategy in probability



theory is the so-called Lindeberg method to replace a non-Gaussian variable by a Gaussian one at each step, and to study the stability of the concerned quantity under such a swapping procedure. A successful use of Lindeberg method to RMT in the recent work of Tao and Vu [30] helped to extend a lot of results on local eigenvalue statistics from GUE to general Wigner matrices. However, Tao and Vu's strategy in [30] requires a detailed analysis on the spectral dynamics of the matrices in the sense that the accurate estimates of the derivatives of the eigenvalues w.r.t the matrix elements are needed.

Later on, Erdős, Yau and Yin proposed another swapping strategy to derive the bulk universality of local statistics in [17]. They studied the stability of the Green function instead of eigenvalues under every swapping step. Such a strategy is based on the elementary resolvent expansion formula (see (2.10)) and turns out to be relatively simpler for certain problems. Very recently, Tao and Vu used a similar swapping strategy on the Green function to derive the CLT for the log-determinant and a sharp concentration of counting functions for Wigner matrices, see [32] and [33]. Note that the objects in [32] and [33] are just two examples of **GLES** with discontinuous test functions (logarithmic and indicator function respectively). It will be clear that one major technical difficulty in our problem is to derive a CLT for **GLES** with the test function continuous but maybe non-differentiable at a few points. Such an ill behaviour in smoothness leads us to pursue the idea in [32] and [33] to study the **GLES** with a class of non-smooth test functions. In Section 5, we will establish a comparison theorem for linear eigenvalue statistics, based on the Helffer-Sjöstrand formula and resolvent expansion. For this purpose, we state some related notions and tools in the remaining part of this section.

The Stieltjes transform of a probability measure  $\mu$  can be defined for all complex number  $z \in \mathbb{C} \setminus \mathbb{R}$  as

$$s^\mu(z) = \int \frac{1}{x - z} \mu(dx).$$

Thus for the ESD  $F_n(x)$  we have

$$s^{M_n}(z) =: s^{F_n}(z) = \frac{1}{n} \sum_{l=1}^n \frac{1}{\lambda_l(M_n) - z} = \frac{1}{n} \text{Tr}(M_n - zI_n)^{-1}.$$

And we also denote the resolvent of  $M_n$  by

$$R^{M_n}(z) = (M_n - zI_n)^{-1},$$

Thus  $s^{M_n}(z) = \frac{1}{n} \text{Tr} R^{M_n}(z)$ . When there is no confusion, we will simplify the symbols  $s^{M_n}(z), R^{M_n}(z)$  by  $s_n(z), R_n(z)$ .

Using the terminology of [32], we say a matrix  $V$  is an elementary matrix if it has one of the following forms

$$V = e_j e_j^*, e_j e_k^* + e_k e_j^*, i e_j e_k^* - i e_k e_j^*$$

with  $1 \leq j \neq k \leq n$ . Here  $e_1, \dots, e_n$  is the standard basis of  $\mathbb{C}^n$ . Let  $M_0$  be an  $n \times n$  Hermitian matrix, and set  $M_t = M_0 + \frac{1}{\sqrt{n}} tV$ . Correspondingly, we denote the resolvent and Stieltjes transform of  $M_t$  by  $R_t(z)$  and  $s_t(z)$  respectively for some complex number  $z = x + iy$  with  $y \neq 0$ . When there is no confusion, we will

simplify the notation  $R_t(z), s_t(z)$  by  $R_t, s_t$ . The notation  $\|A\|_{(\infty,1)}$  for a matrix  $A = (a_{jk})_{j,k=1}^n$  means its  $l^1 \rightarrow l^\infty$  operator norm in the sense that

$$\|A\|_{(\infty,1)} = \sup_{1 \leq j, k \leq n} |a_{jk}|.$$

We conclude this section by the following crucial Taylor expansion for  $s_t$  provided by Tao and Vu.

**Lemma 2.1.** (Proposition 13, [32]) *Suppose that  $x \in \mathbb{R}$ ,  $y > 0$  and  $t \in \mathbb{R}$ . If*

$$|t| \|R_0\|_{(\infty,1)} = o(\sqrt{n}), \quad (2.9)$$

*one has for fixed integer  $k \geq 0$ ,*

$$s_t = s_0 + \sum_{j=1}^k n^{-j/2} c_j t^j + \mathcal{O} \left( n^{-(k+1)/2} |t|^{k+1} \|R_0\|_{(\infty,1)}^{k+1} \min(\|R_0\|_{(\infty,1)}, \frac{1}{ny}) \right)$$

*where the coefficients  $c_j$  are independent of  $t$  and obey the bounds*

$$c_j = \mathcal{O} \left( \|R_0\|_{(\infty,1)}^j \min\{\|R_0\|_{(\infty,1)}, \frac{1}{ny}\} \right)$$

*for all  $1 \leq j \leq k$ .*

Lemma 2.1 is a consequence of the elementary resolvent expansion formula

$$R_t = R_0 + \sum_{j=1}^k \left( -\frac{t}{\sqrt{n}} \right)^j (R_0 V)^j R_0 + \left( -\frac{t}{\sqrt{n}} \right)^{k+1} (R_0 V)^{k+1} R_t. \quad (2.10)$$

We refer to [32] for the details of the proof.

### 3. CLTs FOR GAUSSIAN CASE

First, we truncate the test function so that it is compactly supported and show that such a modification does not alter our results. Set the interval  $\mathcal{U}_\epsilon = [-2-\epsilon, 2+\epsilon]$  with a small constant  $\epsilon < \delta$ . For test function  $f(x)$ , we define the truncated function  $\bar{f}_\epsilon(x) =: \chi_\epsilon(x) f(x)$ . Here  $\chi_\epsilon(x)$  is a smooth cutoff to the region  $|x| \leq 2 + \epsilon$  that equals 1 for  $|x| \leq 2 + \epsilon/2$ . It follows from Lemma 7.4 in Appendix that

$$\mathbb{P}(\mathcal{A}_n[f; u] \neq \mathcal{A}_n[\bar{f}_\epsilon; u]) \leq \mathbb{P}(\max_{1 \leq l \leq n} |\lambda_l| \geq 2 + \epsilon/2) \rightarrow 0$$

as  $n$  goes to infinity. Consequently, without loss of generality, we can thus always assume that  $f$  is compactly supported on the interval  $\mathcal{U}_\epsilon \subset \mathcal{U}$  in this Section.

To prove Theorem 1.1, we start with the type 1 **PLES**. We do the decomposition as follows

$$\mathcal{A}_n[f; u] = \sum_{l=1}^n (f(\lambda_l) - f(u)) \mathbf{1}_{\{\lambda_l \leq u\}} + f(u) N_n(-\infty, u].$$

With the notation defined above, we have

$$\mathcal{A}_n[f; u] = \sum_{l=1}^n f_u(\lambda_l) + f(u) N_n(-\infty, u] = \mathcal{L}_n[f_u] + f(u) N_n(-\infty, u].$$

Observe that  $f_u(x)$  is a continuous function with only one possibly non-differentiable point  $u$ . In order to apply the approach in [22] to treat such a test function, we smooth  $f_u(x)$  in a tiny interval including  $u$ . Set the interval

$$I_n(u) = [u - n^{-1/2-c}, u + n^{-1/2-c}] := [a_n(u), b_n(u)],$$

where  $c$  is a small positive constant. Define the smooth modification of  $f_u(x)$  by

$$g_u(x) =: g_u(n, x) = (f(x) - f(u))\chi_u(n, x)$$

where  $\chi_u(n, x)$  is an  $n$ -dependent smooth cutoff to the region  $x \in (-\infty, b_n(u)]$  that equals 1 for  $x \in (-\infty, a_n(u))$ , and has the property

$$|\frac{d^k}{dx^k}\chi_u(n, x)| = \mathcal{O}(n^{k(1/2+c)}). \quad (3.1)$$

Consequently, one has for  $f \in C^m(\mathcal{U})$ ,

$$|\frac{d^k}{dx^k}g_u(x)| = \mathcal{O}(n^{(k-1)(1/2+c)}), \quad x \in I_n(u), \quad k = 1, \dots, m. \quad (3.2)$$

Using Lemma 7.1 in Appendix one has for some positive constants  $C$

$$N_n(I) \leq Cn|I|$$

with overwhelming probability for any interval  $I$  with length  $|I| \geq n^{-1+c}$ . Together with the trivial fact that

$$\sup_x |f_u(x) - g_u(x)| \leq n^{-1/2-c},$$

we obtain

$$|\mathcal{L}_n[f_u] - \mathcal{L}_n[g_u]| \leq \sup_x |f_u(x) - g_u(x)| N_n(I_n) = o(1) \quad (3.3)$$

with overwhelming probability. Consequently, we have

$$\mathcal{A}_n[f; u] = \mathcal{L}_n[g_u] + f(u)N_n(-\infty, u] + o(1) \quad (3.4)$$

holding with overwhelming probability. Furthermore, we also have

$$\mathbb{E}\mathcal{L}_n[f_u] = \mathbb{E}\mathcal{L}_n[g_u] + o(1). \quad (3.5)$$

As we have mentioned in (1.2), for  $n$ -independent test function  $\varphi \in C_b^1(\mathbb{R})$  with bounded derivative, Lytova and Pastur have proved the CLT. Unfortunately, here our modified test function  $g_u(x)$  is  $n$ -dependent. Thus we can not use Lytova and Pastur's result directly. However, we will show that a slight adjustment of Lytova and Pastur's issue can still lead to the limiting behavior of  $\mathcal{L}_n[g_u]$ . We formulate our conclusion as the following lemma.

**Lemma 3.1.** *If  $M_n$  is drawn from GUE, then for  $f \in C^3(\mathcal{U})$  and compactly supported on  $\mathcal{U}_\epsilon$ , one has*

$$\mathcal{L}_n[g_u] - \mathbb{E}\mathcal{L}_n[g_u] \xrightarrow{d} N(0, V_{GUE}[f_u]).$$

*Proof.* Since we follow the argument of Lytova and Pastur in [22] with only some minor changes, we sketch the proof below. Firstly, we present here some notation and known results laid out in [22]. Let

$$\begin{aligned} U(t) &= e^{itM_n}, & u_n(t) &= \text{Tr}U(t), \\ e_n(x) &= e^{ix\mathcal{L}_n^\circ[g_u]}, & Y_n(x, t) &= \mathbb{E}\{u_n^\circ(t)e_n(x)\}. \end{aligned}$$

The basic idea of [22] is to use the characteristic function to derive a CLT. Set

$$Z_n(x) =: \mathbb{E}e^{ix\mathcal{L}_n^\circ[g_u]} = \mathbb{E}e_n(x).$$

Thus it suffices to show that for any  $x \in \mathbb{R}$

$$\lim_{n \rightarrow \infty} Z_n(x) = Z(x),$$

where

$$Z(x) = \exp\{-x^2 V_{GUE}[f_u]/2\}.$$

Note the relations

$$Z(x) = 1 - V_{GUE}[f_u] \int_0^x y Z(y) dy$$

and

$$Z_n(x) = 1 + \int_0^x Z'_n(y) dy.$$

Using the Cauchy-Schwarz inequality and Proposition 2.4 of [22], we see that

$$|Z'_n(x)| = |i\mathbb{E}\{\mathcal{L}_n^\circ[g_u]e^{ix\mathcal{L}_n^\circ[g_u]}\}| \leq \sqrt{2} \sup_{\lambda} |g'_u(\lambda)| = \mathcal{O}(1)$$

Thus by the dominated convergence theorem, it suffices to verify that any convergent subsequences  $\{Z_{n_j}\}$  and  $\{Z'_{n_j}\}$  satisfy

$$\lim_{n_j \rightarrow \infty} Z_{n_j}(x) = Z(x), \quad \lim_{n_j \rightarrow \infty} Z'_{n_j}(x) = -x V_{GUE}[f_u] Z(x). \quad (3.6)$$

If we denote the Fourier transform of a function  $\varphi$  by

$$\widehat{\varphi}(t) = \frac{1}{2\pi} \int e^{-it\lambda} \varphi(\lambda) d\lambda,$$

we have

$$Z'_n(x) = i\mathbb{E}\{\mathcal{L}_n^\circ[g_u]e^{ix\mathcal{L}_n^\circ[g_u]}\} = i \int \widehat{g}_u(t) Y_n(x, t) dt. \quad (3.7)$$

As shown in [22], to prove (3.6) one needs to prove that the sequence  $\{Y_n\}$  is bounded and equicontinuous on any compact subset of  $\{t \geq 0, x \in \mathbb{R}\}$  (the case of  $t \leq 0$  is analogous), and every uniformly convergent on the set subsequence has the same limit  $Y$ . The proofs for boundness and equicontinuity are really the same as those in [22]. In fact, by the estimates in [22], one has

$$\text{Var}\{u_n(t)\} \leq 2t^2, \quad \text{Var}\{u'_n(t)\} \leq 2(1 + 2t^2), \quad |Y_n(x, t)| \leq \sqrt{2}|t|, \quad (3.8)$$

$$\left| \frac{\partial}{\partial t} Y_n(x, t) \right| \leq \sqrt{2}(1 + 2t^2)^{1/2}, \quad \left| \frac{\partial}{\partial x} Y_n(x, t) \right| \leq 2|t| \sup_{\lambda \in \mathbb{R}} |g'_u(\lambda)| \leq Ct. \quad (3.9)$$

Thus the main task is to show that any uniformly convergent subsequence of  $\{Y_n\}$  has the same limit  $Y$ , and determine the limit. A detailed estimation is presented for Gaussian Orthogonal Ensemble (GOE) in [22]. It is easy to adjust the discussion to GUE case. Applying the calculation procedure of [22] to GUE one can get

$$\begin{aligned} Y_n(x, t) &= -n^{-1} \int_0^t dt_1 \int_0^{t_1} \mathbb{E}\{u_n(t_2 - t_1)u_n(t_2)e_n^\circ(x)\}dt_2 \\ &\quad -x \int_0^t \mathbb{E}\{e_n(x)n^{-1}TrU(t_1)g'_u(M)\}dt_1. \end{aligned}$$

The above equation is just analogous to the corresponding one of the GOE case stated in [22]. Such a representation is a consequence of the integration by parts formula of the Gaussian variables. We refer to [22] for detail. The above equation can be rewritten as

$$\begin{aligned} Y_n(x, t) &+ \int_0^t dt_1 \int_0^{t_1} \bar{v}_n(t_1 - t_2)Y_n(x, t_2)dt_2 \\ &= xZ_n(x)A_n(t) + r_n(x, t), \end{aligned} \quad (3.10)$$

where

$$\begin{aligned} \bar{v}_n(t) &= n^{-1}\mathbb{E}u_n(t), \\ A_n(t) &= -\int_0^t \mathbb{E}\{n^{-1}TrU(t_1)g'_u(M)\}dt_1 \end{aligned}$$

and

$$\begin{aligned} r_n(x, t) &= -n^{-1} \int_0^t dt_1 \int_0^{t_1} \mathbb{E}\{u_n^\circ(t_1 - t_2)u_n^\circ(t_2)e_n^\circ(x)\}dt_2 \\ &\quad -ixn^{-1} \int_0^t dt_1 \int_0^{t_1} t_2\widehat{g}_u(t_2)\mathbb{E}\{u_n(t_1 + t_2)e_n^\circ(x)\}dt_2. \end{aligned} \quad (3.11)$$

By the boundness of  $Y_n(x, t)$ , the first inequality of (3.8) and the Cauchy-Schwarz inequality, one immediately gets that the first term in the expression of  $r_n(x, t)$  is negligible. Now we show that the second term is also  $o(1)$  uniformly in any compact subset of  $\{t \geq 0, x \in \mathbb{R}\}$ . It suffices to prove

$$\int_0^t dt_1 \int_0^{t_1} t_2\widehat{g}_u(t_2)\mathbb{E}\{u_n(t_1 + t_2)e_n^\circ(x)\}dt_2 = o(n). \quad (3.12)$$

In view of (3.8),

$$\begin{aligned} |\mathbb{E}\{u_n(t_1 + t_2)e_n^\circ(x)\}| &= |\mathbb{E}\{u_n^\circ(t_1 + t_2)e_n(x)\}| \\ &\leq Var^{1/2}\{u_n(t_1 + t_2)\} \leq \sqrt{2}|t_1 + t_2|. \end{aligned}$$

Hence it suffices to show that

$$\int (1 + |t|^2)|\widehat{g}_u(t)|dt = o(n).$$

Apparently, we can show for some positive constant  $C$

$$\int_{|t| \geq C} (1 + |t|^2)|\widehat{g}_u(t)|dt = o(n)$$

instead. Note that

$$\begin{aligned}
& \int_{|t| \geq C} (1 + |t|^2) |\widehat{g_u}(t)| dt = \int_{|t| \geq C} \frac{1 + t^2}{|t|^3} |\widehat{g_u^{(3)}}(t)| dt \\
& \leq \left( \int_{|t| \geq C} \left( \frac{1 + t^2}{|t|^3} \right)^2 dt \right)^{1/2} \left( \int |\widehat{g_u^{(3)}}(t)|^2 dt \right)^{1/2} \\
& \leq C \left( \int_{\mathcal{U}_\epsilon} |g_u^{(3)}(x)|^2 dx \right)^{1/2} \\
& \leq C n^{\frac{3}{4} + \frac{3}{2}c},
\end{aligned}$$

where  $g_u^{(3)}(x) = \frac{d^3}{dx^3} g_u(x)$  and  $\widehat{g_u^{(3)}}(t)$  is its Fourier transform. In the above second inequality, we have used the Plancherel's Theorem, and in the last step, we used the bound (3.2). Thus if we choose  $c$  sufficiently small, we can get (3.12). Consequently, we have

$$\lim_{n \rightarrow \infty} r_n(x, t) = 0$$

uniformly on any compact subset of  $\{t \geq 0, x \in \mathbb{R}\}$ .

Moreover, it is not difficult to derive that on any finite interval of  $\mathbb{R}$ ,  $\{\bar{v}_n\}$  and  $\{A_n\}$  converge uniformly to

$$v(t) = \frac{1}{2\pi} \int_{-2}^2 e^{it\lambda} \sqrt{4 - \lambda^2} d\lambda \quad (3.13)$$

and

$$A(t) = -\frac{1}{2\pi} \int_0^t \int_{-2}^u e^{it_1\lambda} f'(\lambda) \sqrt{4 - \lambda^2} d\lambda. \quad (3.14)$$

In fact, the convergence of  $\bar{v}_n$  is a direct consequence of (1.1). For the convergence of  $A_n(t)$ , one can use the convergence rate for ESD of GUE as

$$\sup_x |\mathbb{E}F_n(x) - F_{sc}(x)| \leq Cn^{-1}, \quad (3.15)$$

which was proved by Götze and Tikhomirov in [19]. Note that

$$A_n(t) = - \int_0^t dt_1 \int e^{it_1x} g'_u(x) d\mathbb{E}F_n(x).$$

Using (3.15), by integration by parts, one can easily get

$$A_n(t) = - \int_0^t dt_1 \int e^{it_1x} g'_u(x) dF_{sc}(x) + \mathcal{O}(n^{-1}).$$

Then it is easy to see the right hand side of the above equation tends to  $A(t)$  as  $n$  goes to infinity.

Then by a routine analysis on the limiting equation of (3.10) as that in [22], one can get that  $Y_n(x, t)$  converges to

$$Y(x, t) = \frac{ixZ(x)}{2\pi^2} \int_{-2}^u \int_{-2}^2 \frac{\sqrt{4 - \lambda^2}}{\sqrt{4 - \mu^2}} \frac{e^{it\lambda} - e^{it\mu}}{\lambda - \mu} f'(\lambda) d\lambda d\mu$$

uniformly on any compact subset of  $\{t \geq 0, x \in \mathbb{R}\}$

Note that  $\hat{g}_u(t)$  converges to  $\hat{f}_u(t)$  uniformly in  $t$ . Thus by (3.7) one can get for every convergence subsequence  $\{Z_{n_l}\}_{l \geq 1}$  there exists

$$\lim_{n_l \rightarrow \infty} Z'_{n_l}(x) = -\frac{xZ(x)}{2\pi^2} \int_{-2}^u \int_{-2}^2 \frac{\sqrt{4-\lambda^2}}{\sqrt{4-\mu^2}} \frac{f_u(\lambda) - f_u(\mu)}{\lambda - \mu} f'(\lambda) d\lambda d\mu.$$

By the fact that for  $\lambda \in (-\infty, u)$

$$f'(\lambda)(f_u(\lambda) - f_u(\mu)) = \frac{1}{2} \frac{\partial}{\partial \lambda} (f_u(\lambda) - f_u(\mu))^2,$$

we can use integration by parts to get

$$\begin{aligned} \lim_{n_l \rightarrow \infty} Z'_{n_l}(x) &= -\frac{xZ(x)}{4\pi^2} \int_{-2}^u \int_{-2}^2 \left( \frac{f_u(\lambda) - f_u(\mu)}{\lambda - \mu} \right)^2 \frac{(4 - \lambda\mu)}{\sqrt{4-\lambda^2}\sqrt{4-\mu^2}} d\lambda d\mu \\ &\quad - \frac{xZ(x)}{4\pi^2} \int_{-2}^2 f_u^2(\mu) \frac{\sqrt{4-u^2}}{u-\mu} \frac{1}{\sqrt{4-\mu^2}} d\mu \\ &= -\frac{xZ(x)}{4\pi^2} \int_{-2}^2 \int_{-2}^2 \left( \frac{f_u(\lambda) - f_u(\mu)}{\lambda - \mu} \right)^2 \frac{(4 - \lambda\mu)}{\sqrt{4-\lambda^2}\sqrt{4-\mu^2}} d\lambda d\mu. \end{aligned}$$

Thus we conclude the proof of Lemma 3.1.  $\square$

With the aid of Lemma 3.1, we can now prove Theorem 1.1.

*Proof of Theorem 1.1.* We begin with the CLT for the counting function of eigenvalues of complex Wigner matrices, whose proof can be found in the recent work of Dallaporta and Vu [10].

**Lemma 3.2** ([10]). *If  $M_n$  is a complex Wigner matrix satisfying Condition  $\mathbf{C}_0$  and matches to GUE to the fourth order off the diagonal and the second order on the diagonal, one has*

$$\frac{N_n(-\infty, u] - \mathbb{E}N_n(-\infty, u]}{\sqrt{\frac{1}{2\pi^2} \log n}} \xrightarrow{d} N(0, 1), \quad (3.16)$$

where

$$\mathbb{E}N_n(-\infty, u] = n \frac{1}{2\pi} \int_{-2}^u \sqrt{4-x^2} dx + o(1). \quad (3.17)$$

For  $f \in C^1(\mathbb{R})$ , we also have  $g_u \in C^1(\mathbb{R})$ . By Proposition 2.4 of [22], we have for GUE

$$\text{Var} \mathcal{L}_n[g_u] \leq 2 \left( \sup_{x \in \mathbb{R}} |g'_u(x)| \right)^2 \leq C.$$

Clearly,

$$\frac{\mathcal{L}_n^\circ[g_u]}{\sqrt{\log n}} \xrightarrow{P} 0 \quad (3.18)$$

as  $n$  goes to infinity. Combining (3.4), (3.5), (3.16) and (3.18), we immediately get that when  $f \in C^1(\mathbb{R})$  with bounded derivative and  $f(u) \neq 0$ ,

$$\frac{\mathcal{A}_n[f; u] - \mathbb{E}\mathcal{L}_n[f_u] - f(u)\mathbb{E}N_n(-\infty, u]}{\sqrt{\frac{f^2(u)}{2\pi^2} \log n}} \xrightarrow{d} N(0, 1).$$

To prove (i) of Theorem 1.1, it remains to show for  $f \in C^1(\mathcal{U})$  compactly supported on  $\mathcal{U}_\epsilon$ ,

$$\mathbb{E}\mathcal{L}_n[f_u] = n \frac{1}{2\pi} \int_{-2}^u f_u(x) \sqrt{4-x^2} dx + \mathcal{O}(1). \quad (3.19)$$

Observe that by (3.15),

$$\begin{aligned} |\mathbb{E}\mathcal{L}_n[f_u] - n \int_{-2}^u f_u(x) dF_{sc}(x)| &= n \left| \int_{-2-\epsilon}^u f_u(x) d\mathbb{E}F_n(x) - \int_{-2}^u f_u(x) dF_{sc}(x) \right| \\ &= n \left| \int_{-2-\epsilon}^u f'_u(x) (\mathbb{E}F_n(x) - F_{sc}(x)) dx \right| \\ &\leq n \int_{-2-\epsilon}^u |f'_u(x)| |\mathbb{E}F_n(x) - F_{sc}(x)| dx \\ &= \mathcal{O}(1), \end{aligned}$$

which implies (3.19). Thus we complete the proof of (i).

Now we turn to the case where  $f(u) = 0$ . Since  $f \in C^4(\mathcal{U})$ , by Lemma 3.1, together with (3.4) and (3.5) we can easily obtain

$$\frac{\mathcal{A}_n[f; u] - \mathbb{E}\mathcal{L}_n[f_u]}{\sqrt{V_{GUE}[f_u]}} \xrightarrow{d} N(0, 1).$$

Thus to prove (ii) of Theorem 1.1, it suffices to show for  $f \in C^4(\mathcal{U})$  and compactly supported on  $\mathcal{U}_\epsilon$  the more accurate estimate

$$\mathbb{E}\mathcal{L}_n[f_u] = n \frac{1}{2\pi} \int_{-2}^u f_u(x) \sqrt{4-x^2} dx + o(1). \quad (3.20)$$

To show (3.20), we define two smooth cutoff functions  $\chi_1(x)$  and  $\chi_2(x)$ . Let  $\chi_1(x)$  be a smooth cutoff function which is equal to 1 for  $x \geq -1 + u/2$  and 0 for  $x \leq -3/2 + u/4$ , such that  $\chi_1^{(k)}(x) \leq C$  holds for some positive constant  $C$  and  $k = 0, \dots, 4$ . Let  $\chi_2(x) = 1 - \chi_1(x)$ . Now we decompose  $f_u(x)$  as

$$f_u(x) = \chi_1(x)f_u(x) + \chi_2(x)f_u(x).$$

Observe that  $\chi_2(x)f_u(x) \in C^4(\mathcal{U})$ . It has been proved in Bai, Wang and Zhou [3] that for  $C^4(\mathcal{U})$  function supported on  $\mathcal{U}_\epsilon$ , one has

$$\mathbb{E}\mathcal{L}_n[\chi_2(x)f_u(x)] = n \int \chi_2(x)f_u(x) dF_{sc}(x) + o(1). \quad (3.21)$$

For  $\mathbb{E}\mathcal{L}_n[\chi_1(x)f_u(x)]$ , we use the following asymptotic formula proved in Ercolani and McLaughlin [14],

$$\mathcal{K}_n(x, x) = \frac{n}{2\pi} \sqrt{4-x^2} + \frac{1}{4\pi} \left( \frac{1}{x-2} - \frac{1}{x+2} \right) \cos \left[ \frac{n}{2\pi} \int_x^2 \sqrt{4-y^2} dy \right] + \mathcal{O}(n^{-1})$$



for  $x \in [-2 + \delta, 2 - \delta]$  with any fixed  $\delta > 0$ . Thus by using (2.2) one has

$$\mathbb{E}\mathcal{L}_n[\chi_1(x)f_u(x)] = n \int \chi_1(x)f_u(x)dF_{sc}(x) + \varepsilon_n,$$

where

$$\varepsilon_n = \int_{-\frac{3}{2}+\frac{u}{4}}^u \chi_1(x)f_u(x)\frac{1}{4\pi}\left(\frac{1}{x-2} - \frac{1}{x+2}\right)\cos\left[\frac{n}{2\pi}\int_x^2\sqrt{4-y^2}dy\right]dx + \mathcal{O}(n^{-1}).$$

By integration by parts, we can easily get that

$$\varepsilon_n = \mathcal{O}(n^{-1}). \quad (3.22)$$

Thus combining (3.21) with (3.22) we can show (3.20).

Now we prove (iii) of Theorem 1.1. By definition,  $\gamma_{k/n}$  is the  $k$ -th  $n$ -quantile of the semicircle law, i.e.

$$\frac{1}{2\pi} \int_{-2}^{\gamma_{k/n}} \sqrt{4-x^2}dx = \frac{k}{n}.$$

We decompose  $\mathcal{B}_n[f; k]$  as

$$\mathcal{B}_n[f; k] = \sum_{l=1}^k (f(\lambda_l) - f(\gamma_{k/n})) + kf(\gamma_{k/n}).$$

In order to avoid working on the ordered eigenvalues, we introduce a proxy of  $\mathcal{B}_n[f; k]$  as

$$\begin{aligned} \widehat{\mathcal{B}}_n[f; k] &= \sum_{l=1}^n (f(\lambda_l) - f(\gamma_{k/n}))\mathbf{1}_{\{\lambda_l \leq \gamma_{k/n}\}} + kf(\gamma_{k/n}) \\ &= \sum_{l=1}^{nF_n(\gamma_{k/n})} (f(\lambda_l) - f(\gamma_{k/n})) + kf(\gamma_{k/n}). \end{aligned}$$

Let  $a = |nF_n(\gamma_{k/n}) - k|$ . By using the rigidity property in Lemma 7.2, one has with overwhelming probability

$$a = |nF_n(\gamma_{k/n}) - nF_{sc}(\gamma_{k/n})| \leq (\log n)^{C \log \log n}.$$

Furthermore, we also have

$$|\mathcal{B}_n[f; k] - \widehat{\mathcal{B}}_n[f; k]| \leq a \max_{l \in \{k-a, k+a\}} |f(\lambda_l) - f(\gamma_{k/n})| = \mathcal{O}\left(\frac{(\log n)^{C \log \log n}}{n}\right) \quad (3.23)$$

with overwhelming probability. Therefore, we only have to prove the central limit theorem for  $\widehat{\mathcal{B}}_n[f; k]$ .

Observe that

$$\widehat{\mathcal{B}}_n[f; k] = \sum_{l=1}^n f_{\gamma_{k/n}}(\lambda_l) + kf(\gamma_{k/n}) = \mathcal{L}_n[f_{\gamma_{k/n}}] + kf(\gamma_{k/n}). \quad (3.24)$$

Moreover, since we assume  $f$  is compactly supported on  $\mathcal{U}_\epsilon$ , we have

$$\mathbb{E}\mathcal{L}_n[f_{\gamma_{k/n}}] = m[f; \gamma_{k/n}] + o(1); \quad \mathbb{E}\mathcal{L}_n[f_{\gamma_y}] = m[f; \gamma_y] + o(1)$$

as shown in (3.20). Thus to prove (iii), it suffices to show as  $n \rightarrow \infty$

$$|\mathbb{E}e^{-ix\mathcal{L}_n^\circ[f_{\gamma_{k/n}}]} - \mathbb{E}e^{-ix\mathcal{L}_n^\circ[f_{\gamma_y}]}| \rightarrow 0 \quad (3.25)$$

for any fixed  $x$ . To see (3.25), we note that

$$\begin{aligned} & |\mathbb{E}e^{-ix\mathcal{L}_n^\circ[f_{\gamma_{k/n}}]} - \mathbb{E}e^{-ix\mathcal{L}_n^\circ[f_{\gamma_y}]}| \\ & \leq |x|\mathbb{E}|\mathcal{L}_n^\circ[f_{\gamma_{k/n}}] - \mathcal{L}_n^\circ[f_{\gamma_y}]| \\ & \leq |x|Var^{1/2}\{\mathcal{L}_n[f_{\gamma_{k/n}} - f_{\gamma_y}]\}. \end{aligned}$$

Thus it remains to verify

$$Var\mathcal{L}_n[f_{\gamma_{k/n}} - f_{\gamma_y}] \rightarrow 0. \quad (3.26)$$

To show this, we will rely on the following lemma whose proof will be postponed to the end of this section. It will be clear that the following lemma is also crucial to our proof of tightness for  $\mathcal{S}_n^\circ[f; t]$  in the next section.

**Lemma 3.3.** *Suppose that  $\varphi$  is a Lipschitz function on  $\mathbb{R}$  with Lipschitz constant  $L$ . Moreover, we assume that there exists an interval  $I = [a, b] \subset [-2 + \delta, 2 - \delta]$  ( $a, b$  may be  $n$ -dependent) such that*

$$\varphi(\lambda) = \varphi(a), \quad \lambda \leq a; \quad \varphi(\lambda) = \varphi(b), \quad \lambda \geq b.$$

Then for GUE, we have

$$Var\mathcal{L}_n[\varphi] \leq C(b-a)^2(|\log(b-a)| + 1) + \mathcal{O}(e^{-cn})$$

with some positive constants  $C =: C(L)$  and  $c$  independent of  $a$  and  $b$ .

Now we proceed to the proof of (iii) of Theorem 1.1. For convenience, we assume  $k/n \geq y$ . The opposite case is just analogous. Let  $\varphi = f_{\gamma_{k/n}} - f_{\gamma_y}$ . By definition, we note that  $\varphi(\lambda)$  equals to  $f(\gamma_y) - f(\gamma_{k/n})$  for  $\lambda \leq \gamma_y$  and 0 for  $\lambda \geq \gamma_{k/n}$ . Thus by Lemma 3.3 and the assumption that  $k/n \rightarrow y$ , we have (3.26). Thus (3.25) holds. So we conclude the proof of (iii) by using (3.25) and (ii) of Theorem 1.1.  $\square$

*Proof of Lemma 3.3.* From the determinantal structure of the j.p.d. of eigenvalues for GUE, with (2.3) one has

$$Var\mathcal{L}_n[\varphi] = \int_{\mathbb{R}} \int_{\mathbb{R}} \left| \frac{\varphi(\lambda) - \varphi(\mu)}{\lambda - \mu} \right|^2 \mathcal{V}_n(\lambda, \mu) d\lambda d\mu,$$

where

$$\begin{aligned} \mathcal{V}_n(\lambda, \mu) &= \frac{1}{2} \mathcal{K}_n^2(\lambda, \mu) (\lambda - \mu)^2 \\ &= \frac{1}{2} (\sqrt{n}\psi_n(\sqrt{n}x)\psi_{n-1}(\sqrt{n}y) - \sqrt{n}\psi_{n-1}(\sqrt{n}x)\psi_n(\sqrt{n}y))^2. \end{aligned}$$

By assumption, we can split the integral into four parts

$$Var\mathcal{L}_n[\varphi] = V_1 + V_2 + V_3 + V_4$$

where

$$V_1 =: \int_a^b \int_a^b \left| \frac{\varphi(\lambda) - \varphi(\mu)}{\lambda - \mu} \right|^2 \mathcal{V}_n(\lambda, \mu) d\lambda d\mu$$

$$\begin{aligned}
V_2 &=: 2 \int_a^b \int_{-\infty}^a \left| \frac{\varphi(\lambda) - \varphi(\mu)}{\lambda - \mu} \right|^2 \mathcal{V}_n(\lambda, \mu) d\lambda d\mu \\
V_3 &=: 2 \int_a^b \int_b^\infty \left| \frac{\varphi(\lambda) - \varphi(\mu)}{\lambda - \mu} \right|^2 \mathcal{V}_n(\lambda, \mu) d\lambda d\mu \\
V_4 &=: 2 \int_{-\infty}^a \int_b^\infty \left| \frac{\varphi(\lambda) - \varphi(\mu)}{\lambda - \mu} \right|^2 \mathcal{V}_n(\lambda, \mu) d\lambda d\mu.
\end{aligned}$$

Note that  $\varphi$  is Lipschitz. And it is well known that  $\mathcal{V}_n(x, y)$  is bounded in  $\mathbb{R}^2$ . Moreover, by (2.5) one sees that  $\mathcal{V}_n(x, y)$  is exponentially decaying in  $x, y$  as  $|x|$  or  $|y|$  larger than  $2 + \delta$ . Thus we immediately get that

$$\begin{aligned}
V_1 &\leq C(b-a)^2, \\
V_2 &=: 2 \int_a^b \int_{-2-\delta}^a \left| \frac{\varphi(\lambda) - \varphi(\mu)}{\lambda - \mu} \right|^2 \mathcal{V}_n(\lambda, \mu) d\lambda d\mu + \mathcal{O}(e^{-cn}), \\
V_3 &=: 2 \int_a^b \int_b^{2+\delta} \left| \frac{\varphi(\lambda) - \varphi(\mu)}{\lambda - \mu} \right|^2 \mathcal{V}_n(\lambda, \mu) d\lambda d\mu + \mathcal{O}(e^{-cn}), \\
V_4 &=: 2 \int_{-2-\delta}^a \int_b^{2+\delta} \left| \frac{\varphi(\lambda) - \varphi(\mu)}{\lambda - \mu} \right|^2 \mathcal{V}_n(\lambda, \mu) d\lambda d\mu + \mathcal{O}(e^{-cn}).
\end{aligned}$$

Now we estimate  $V_2$ . Note that

$$\begin{aligned}
V_2 &= 2 \int_a^b \int_{2a-b}^a \left| \frac{\varphi(\lambda) - \varphi(\mu)}{\lambda - \mu} \right|^2 \mathcal{V}_n(\lambda, \mu) d\lambda d\mu \\
&+ 2 \int_a^b \int_{-2-\delta}^{2a-b} \left| \frac{\varphi(\lambda) - \varphi(\mu)}{\lambda - \mu} \right|^2 \mathcal{V}_n(\lambda, \mu) d\lambda d\mu + \mathcal{O}(e^{-cn}).
\end{aligned}$$

Observe that the first term on the right hand side of the above equality can be bounded by  $C(b-a)^2$ , and the second term can be bounded as

$$\begin{aligned}
&\int_a^b \int_{-2-\delta}^{2a-b} \left| \frac{\varphi(\lambda) - \varphi(\mu)}{\lambda - \mu} \right|^2 \mathcal{V}_n(\lambda, \mu) d\lambda d\mu \\
&\leq C(b-a)^2 \int_a^b \int_{-2-\delta}^{2a-b} \left| \frac{1}{\lambda - \mu} \right|^2 d\lambda d\mu \\
&\leq C(b-a)^2 (|\log(b-a)| + 1).
\end{aligned}$$

Thus we have

$$V_2 \leq C(b-a)^2 (|\log(b-a)| + 1) + \mathcal{O}(e^{-cn}).$$

Analogously, one can also get that

$$V_3, V_4 \leq C(b-a)^2 (|\log(b-a)| + 1) + \mathcal{O}(e^{-cn}).$$

Thus we conclude the proof.  $\square$

## 4. PARTIAL SUM PROCESS FOR GUE

In this section, we provide the proof of Theorem 1.4. Thus we have to verify the finite dimensional convergence and the tightness of the sequence  $\{\mathcal{S}_n^\circ[f; t]; t \in [\delta, 1 - \delta]\}$ .

At first, we extend the discussion in the last section to show that the finite dimensional convergence of the process  $\{\mathcal{S}_n^\circ[f; t]; t \in [\delta, 1 - \delta]\}$ . We formulate the result as the following lemma.

**Lemma 4.1.** *Under the assumptions of Theorem 1.4, for any fixed positive integer  $r$  and points  $t_1, \dots, t_r \in [\delta, 1 - \delta]$ , and for any fixed numbers  $\alpha_1, \dots, \alpha_r \in \mathbb{R}$ , we have*

$$\sum_{l=1}^r \alpha_l \mathcal{S}_n^\circ[f; t_l] \xrightarrow{d} \sum_{l=1}^r \alpha_l \mathcal{S}[f; t_l].$$

*Proof.* Below we set  $k_l = \lfloor nt_l \rfloor$  and  $u(t_l) = k_l/n$ . At first, we claim that for integer  $l \in [\delta n, (1 - \delta)n]$ ,

$$\text{Var}\{f(\lambda_l)\} \leq C \frac{\log n}{n^2} \quad (4.1)$$

for  $f$  obeying the assumptions in Theorem 1.4. Here the constant  $C$  depends only on  $\delta$  and the function  $f$ . To show (4.1), we recall the bump function  $\chi_\epsilon$  and the corresponding truncated function  $\tilde{f}_\epsilon(x)$  defined in Section 3. Then we put

$$\tilde{f}_\epsilon(x) = f(x) - \bar{f}_\epsilon(x),$$

which vanishes when  $|x| \leq 2 + \epsilon/2$ . Therefore, one has

$$\text{Var}\{f(\lambda_l)\} \leq 2\text{Var}\{\tilde{f}_\epsilon(\lambda_l)\} + 2\mathbb{E}(\tilde{f}_\epsilon(\lambda_l))^2.$$

It follows from Lemma 7.5 that

$$\text{Var}\{\tilde{f}_\epsilon(\lambda_l)\} \leq \sup_x |\tilde{f}_\epsilon'(x)|^2 \text{Var}\{\lambda_l\} \leq C \frac{\log n}{n^2}.$$

Besides, by (2.6) and the assumptions on  $f(x)$ , we also have

$$\mathbb{E}(\tilde{f}_\epsilon(\lambda_l))^2 = \mathcal{O}(e^{-cn}) \leq C \frac{\log n}{n^2}$$

for sufficiently large  $n$ . Thus we have (4.1). Consequently, we have

$$\sum_{l=1}^r \alpha_l f^\circ(\lambda_{k_l+1}) \xrightarrow{P} 0.$$

Then it remains to show

$$\sum_{l=1}^r \alpha_l B_n^\circ[f; k_l] \xrightarrow{d} \sum_{l=1}^r \alpha_l \mathcal{S}[f; t_l].$$

Using (3.23) and (3.24), it suffices to prove

$$\sum_{l=1}^r \alpha_l \mathcal{L}_n^\circ[f_{\gamma_{u(t_l)}}] \xrightarrow{d} \sum_{l=1}^r \alpha_l \mathcal{S}[f; t_l].$$

By the fact that  $u(t_l) \rightarrow t_l$  and an routine discussion as that for (3.25), we can reduce the problem to show that

$$\sum_{l=1}^r \alpha_l \mathcal{L}_n^\circ[f_{\gamma_{t_l}}] \xrightarrow{d} \sum_{l=1}^r \alpha_l \mathcal{S}[f; t_l].$$

Note that

$$\sum_{l=1}^r \alpha_l \mathcal{L}_n^\circ[f_{\gamma_{t_l}}] = \mathcal{L}_n^\circ\left[\sum_{l=1}^r \alpha_l f_{\gamma_{t_l}}\right].$$

Observe that  $\sum_{l=1}^r \alpha_l f_{\gamma_{t_l}}$  is a continuous function with  $r$  possibly non differentiable points  $t_1, \dots, t_r$ . Now we choose  $r$  interval  $I_1, \dots, I_r$  containing  $\gamma_{t_1}, \dots, \gamma_{t_r}$  respectively with lengths  $|I_l| \leq n^{-1/2-c}$  for some small positive number  $c$  and all  $l = 1, \dots, r$ . We define a smooth modification function  $g_{t_1, \dots, t_r}$  which coincides with  $\sum_{l=1}^r \alpha_l f_{\gamma_{t_l}}$  on  $\mathbb{R} \setminus \cup_{l=1}^r I_l$  and obeys the condition

$$\left| \frac{d^{k+1}}{dx^{k+1}} g_{t_1, \dots, t_r}(x) \right| = \mathcal{O}(n^{k(1/2+c)}), k = 0, 1, 2.$$

By a similar relation to (3.3), we only have to prove

$$\mathcal{L}_n^\circ[g_{t_1, \dots, t_r}] \xrightarrow{d} \sum_{l=1}^r \alpha_l \mathcal{S}[f; t_l]. \quad (4.2)$$

The proof of (4.2) is easy to carry out by using Lytova and Pastur's method again as that in Section 3. Thus we can finally get that  $\sum_{l=1}^r \alpha_l \mathcal{S}_n^\circ[f; t_l]$  weakly converges to a Gaussian distribution with mean zero and variance  $V_{GUE}[\sum_{l=1}^r \alpha_l f_{\gamma_{t_l}}]$ . Consequently, we can conclude the proof of Lemma 4.1.  $\square$

It remains to show that the sequence  $\{\mathcal{S}_n^\circ[f; t]; t \in [\delta, 1 - \delta]\}$  is tight. To this end, we will use Theorem 12.3 (p. 95) of Billingsley [6]. According to this theorem, we need to verify

(i): *Tightness at any point in  $[\delta, 1 - \delta]$ .*

(ii): *For arbitrary  $s, t \in [\delta, 1 - \delta]$  and  $n$  sufficiently large*

$$\mathbb{E}|\mathcal{S}_n^\circ[f; t] - \mathcal{S}_n^\circ[f; s]|^2 \leq C|t - s|^\alpha$$

for some constant  $C > 0$ , and  $\alpha > 1$  which are independent of  $t, s$ .

Note that (i) is obvious. Thus it suffices to show (ii). Set  $\eta_n = \log^{-L} n$  for some constant  $L$  large enough. Without loss of generality, we always assume that  $s \leq t$  below. We separate the issue into three cases:  $0 \leq t - s \leq n^{-1}$ ,  $n^{-1} \leq t - s \leq \eta_n$  and  $t - s \geq \eta_n$ .

For  $t - s \leq n^{-1}$ , one has  $\lfloor nt \rfloor = \lfloor ns \rfloor$  or  $\lfloor nt \rfloor = \lfloor ns \rfloor + 1$ . When  $\lfloor nt \rfloor = \lfloor ns \rfloor$ , we have

$$\mathbb{E}|\mathcal{S}_n^\circ[f; t] - \mathcal{S}_n^\circ[f; s]|^2 = \mathbb{E}|n(t - s)f^\circ(\lambda_{\lfloor ns \rfloor + 1})|^2 \leq C|t - s|^2 \log n.$$

In the last step above we used the estimation (4.1). And the positive constant  $C$  only depends on  $\delta$  and the test function  $f$  (Such a dependence will not be mentioned repeatedly below for simplicity). When  $\lfloor nt \rfloor = \lfloor ns \rfloor + 1$ , one has

$$\mathbb{E}|\mathcal{S}_n^\circ[f; t] - \mathcal{S}_n^\circ[f; s]|^2 = \mathbb{E}|(nt - \lfloor nt \rfloor)f^\circ(\lambda_{\lfloor nt \rfloor + 1}) + (\lfloor nt \rfloor - ns)f^\circ(\lambda_{\lfloor nt \rfloor})|^2.$$

Note that when  $\lfloor nt \rfloor = \lfloor ns \rfloor + 1$ , one has

$$0 \leq nt - \lfloor nt \rfloor, \lfloor nt \rfloor - ns \leq nt - ns.$$

Consequently, we have

$$\begin{aligned} & \mathbb{E}|\mathcal{S}_n^\circ[f; t] - \mathcal{S}_n^\circ[f; s]|^2 \\ & \leq Cn^2(t-s)^2(\text{Var}\{f(\lambda_{\lfloor nt \rfloor + 1})\} + \text{Var}\{f(\lambda_{\lfloor nt \rfloor})\}) \\ & \leq C|t-s|^2 \log n. \end{aligned}$$

Since  $0 \leq t-s \leq n^{-1}$ , for  $n$  large enough, we always have

$$|t-s|^2 \log n \leq |t-s|^{3/2}.$$

For  $n^{-1} \leq t-s \leq \eta_n$ , one has

$$\begin{aligned} \mathbb{E}|\mathcal{S}_n^\circ[f; t] - \mathcal{S}_n^\circ[f; s]|^2 & \leq 2\mathbb{E}|\mathcal{B}_n^\circ[f, \lfloor nt \rfloor] - \mathcal{B}_n^\circ[f, \lfloor ns \rfloor]|^2 \\ & \quad + 2\mathbb{E}|(nt - \lfloor nt \rfloor)f^\circ(\lambda_{\lfloor nt \rfloor + 1}) - (ns - \lfloor ns \rfloor)f^\circ(\lambda_{\lfloor ns \rfloor + 1})|^2 \\ & \leq 2(\lfloor nt \rfloor - \lfloor ns \rfloor) \sum_{i=\lfloor ns \rfloor}^{\lfloor nt \rfloor} \text{Var}\{f(\lambda_i)\} \\ & \quad + 2\mathbb{E}|(nt - \lfloor nt \rfloor)f^\circ(\lambda_{\lfloor nt \rfloor + 1}) - (ns - \lfloor ns \rfloor)f^\circ(\lambda_{\lfloor ns \rfloor + 1})|^2 \\ & \leq C(t-s)^2 \log n. \end{aligned}$$

In the above second inequality we have used the basic relation

$$\text{Var}\{\xi_1 + \cdots + \xi_m\} \leq m \sum_{l=1}^m \text{Var}\{\xi_l\}.$$

Clearly, when the constant  $L$  in the definition of  $\eta_n$  is chosen to be large enough, for  $n^{-1} \leq t-s \leq \eta_n$  we have

$$(t-s)^2 \log n \leq (t-s)^{3/2}.$$

For the last case  $t-s \geq \eta_n$ , it suffices to show that

$$\mathbb{E}|\mathcal{B}_n^\circ[f, \lfloor nt \rfloor] - \mathcal{B}_n^\circ[f, \lfloor ns \rfloor]|^2 \leq C|t-s|^{3/2}.$$

By (3.23) and (3.24) we need to prove

$$\text{Var}|\mathcal{L}_n[f_{\gamma_{u(t)}} - f_{\gamma_{u(s)}}]|^2 \leq C|t-s|^{3/2}, \quad (4.3)$$

where  $u(t) = \lfloor nt \rfloor/n$ . Note that (4.3) follows from Lemma 3.3 immediately. So Theorem 1.4 follows.

## 5. CLTs FOR WIGNER MATRICES

As shown in Section 3, Lytova and Pastur's original proof in [22] for the GUE case can be easily modified to adapt to our case. However, for more general complex Wigner matrix, higher order derivatives of  $g_u(x)$  will be involved if we proceed to pursue the discussion in [22] (see (3.49) of [22] for instance). But those derivatives in (3.1) will not be small enough for the strategy in [22]. Moreover, the results in [22] do not provide the asymptotic estimation of the expectation.

Motivated by the recent articles [32] and [33], we will establish a comparison theorem for the linear eigenvalue statistics with a certain class of test functions in this section. As an application, we use the comparison theorem to extend Theorem 1.1 to general complex Wigner matrices case. Moreover, such a comparison theorem will also be used in the next section to prove Theorem 1.8.

At first, we define the set of  $n$ -dependent real functions  $\mathcal{F}_n^m$  for some fixed positive integer  $m$ . We say a function  $\varphi \in \mathcal{F}_n^m$  if and only if  $\varphi =: \varphi_n$  satisfies the following assumptions **(a)** and **(b)**.

**(a)**:  $\varphi \in C^4(\mathcal{U})$  is compactly supported on  $\mathcal{U}_\epsilon$  and

$$|\varphi^{(\alpha)}(x)| \leq C, \quad \alpha = 0, 1$$

with some positive constant  $C$  independent of  $n$ .

**(b)**: There exist  $m$  intervals  $I_1, \dots, I_m \in [-2 + \delta, 2 - \delta]$  with length  $|I_l| \leq n^{-1-c_1}$  for some small  $c_1 > 0$  and all  $l = 1, \dots, m$ , such that for  $x \in \mathcal{U} \setminus \cup_{l=1}^m I_l$  one has

$$|\varphi^{(\alpha)}(x)| \leq C, \quad \alpha = 2, 3, 4$$

with some positive constant  $C$  independent of  $n$ .

Our main tool to extend the CLTs from GUE to general Wigner matrices is the following comparison theorem for linear eigenvalue statistics.

**Theorem 5.1.** Let  $M_n = \frac{1}{\sqrt{n}}(w_{jk})_{j,k=1}^n$  and  $M'_n = \frac{1}{\sqrt{n}}(w'_{jk})_{j,k=1}^n$  be two Wigner matrices satisfying Condition **C<sub>0</sub>**. We assume  $M_n$  and  $M'_n$  match to the fourth order off the diagonal and to the second order on the diagonal. Moreover, the magnitudes of  $w_{jk}$  and  $w'_{jk}$  ( $1 \leq j, k \leq n$ ) are uniformly bounded by  $n^{\mathcal{O}(c_0)}$  for some sufficiently small but fixed  $c_0 > 0$ . Let  $G : \mathbb{R} \rightarrow \mathbb{R}$  obey the derivative bounds

$$\left| \frac{d^j G(x)}{dx^j} \right| = \mathcal{O}(n^{c_0}) \quad (5.1)$$

for  $0 \leq j \leq 5$ . If  $\varphi \in \mathcal{F}_n^m$  for some fixed integer  $m$ , then we have

$$\mathbb{E}G(\mathcal{L}_n^M[\varphi]) - \mathbb{E}G(\mathcal{L}_n^{M'}[\varphi]) = \mathcal{O}(n^{-c})$$

for some fixed  $c > 0$ . Here  $\mathcal{L}_n^M[\varphi]$  (resp.  $\mathcal{L}_n^{M'}[\varphi]$ ) represents the linear eigenvalue statistic of  $M_n$  (resp.  $M'_n$ ) with the test function  $\varphi$ .

Using the terminology in [32], we say a statistic  $S(M_n)$  that can depend on  $M_n$  or  $M'_n$  highly insensitive if one has

$$|S(M_n) - S(M'_n)| = \mathcal{O}(n^{-c}).$$

for some fixed  $c > 0$ . Thus Theorem 5.1 asserts that  $\mathbb{E}G(\mathcal{L}_n^M[\varphi])$  is highly insensitive for  $\varphi \in \mathcal{F}_n^m$ . To show this, our strategy is to represent the linear eigenvalue statistics by the Stieltjes transform of the ESD defined in Section 2. Then the Lindeberg swapping argument for the Stieltjes transform which was well developed in recent work such as [18] and [32] can be applied. To this end, we use the following Helffer-Sjöstrand formula.

**Lemma 5.2** (Helffer-Sjöstrand formula). Suppose that  $\varphi : \mathbb{R} \rightarrow \mathbb{R}$  be a  $C^{k+1}(\mathbb{R})$  function with a compact support. Let  $\sigma(y) \in C^\infty(\mathbb{R})$  be a cut off function such that

$\sigma(y) = 1$  for  $|y| \leq 1/2$  and  $\sigma(y) = 0$  for  $|y| \geq 1$  with bounded derivatives. Define the smooth extension  $\tilde{\varphi} : \mathbb{C} \rightarrow \mathbb{C}$  of  $\varphi$  by

$$\tilde{\varphi}(z) := \left( \sum_{j=0}^k \frac{\varphi^{(j)}(x)(iy)^j}{j!} \right) \sigma(y),$$

where  $z = x + iy$ . Then for any self-adjoint operator  $X$ , one has

$$\varphi(X) = \frac{1}{\pi} \int_{\mathbb{R}^2} \frac{\partial \tilde{\varphi}}{\partial \bar{z}} \frac{1}{X - z} dx dy,$$

where

$$\frac{\partial \tilde{\varphi}}{\partial \bar{z}} := \frac{1}{2} \left( \frac{\partial \tilde{\varphi}}{\partial x} + i \frac{\partial \tilde{\varphi}}{\partial y} \right).$$

**Remark 5.3.** We refer to Davies' book [11] for more details on the Helffer-Sjöstrand formula. Moreover, in the literature of RMT, one can also see [16] and [26] for references.

By definition, one can calculate

$$\frac{\partial \tilde{\varphi}}{\partial \bar{z}} = \frac{1}{2} \left( \sum_{j=0}^k \frac{\varphi^{(j)}(x)(iy)^j}{j!} \right) i\sigma'(y) + \frac{1}{2k!} \varphi^{(k+1)}(x)(iy)^k \sigma(y).$$

Below we will restrict to the case of  $m = 1$  for ease of presentation. It will be clear that the proof can be extended straightforward to the case of  $m > 1$  but fixed. For simplicity, we will denote  $I_1$  and  $\mathcal{F}_n^1$  by  $I$  and  $\mathcal{F}_n$  respectively. For  $\varphi \in \mathcal{F}_n$ , we denote  $I =: [a, b]$  with  $a = a_n, b = b_n$  such that  $|b - a| \leq n^{-1-c_1}$ . By the Helffer-Sjöstrand formula we can write

$$\begin{aligned} \mathcal{L}_n[\varphi] &= \frac{1}{\pi} \sum_{l=1}^n \int_{\mathbb{R}^2} \frac{\partial \tilde{\varphi}}{\partial \bar{z}} \frac{1}{\lambda_l - z} dx dy \\ &= \frac{1}{2\pi} \sum_{l=1}^n \int_{\mathbb{R}^2} \left( \sum_{j=0}^3 \frac{\varphi^{(j)}(x)(iy)^j}{j!} \right) i\sigma'(y) \frac{1}{\lambda_l - z} dx dy \\ &\quad + \frac{1}{2\pi} \sum_{l=1}^n \int_{\mathbb{R}^2} \frac{1}{3!} \varphi^{(4)}(x)(iy)^3 \sigma(y) \frac{1}{\lambda_l - z} dx dy \\ &= \frac{1}{2\pi} \sum_{l=1}^n \int_{-2-\epsilon}^{2+\epsilon} \left( \int_{-1}^{-1/2} + \int_{1/2}^1 \right) \left( \sum_{j=0}^3 \frac{\varphi^{(j)}(x)(iy)^j}{j!} \right) i\sigma'(y) \frac{1}{\lambda_l - z} dx dy \\ &\quad + \frac{1}{2\pi} \sum_{l=1}^n \left( \int_b^{2+\epsilon} + \int_{-2-\epsilon}^a \right) \int_{-1}^1 \frac{1}{3!} \varphi^{(4)}(x)(iy)^3 \sigma(y) \frac{1}{\lambda_l - z} dx dy \\ &\quad + \frac{1}{2\pi} \sum_{l=1}^n \int_a^b \int_{-1}^1 \frac{1}{3!} \varphi^{(4)}(x)(iy)^3 \sigma(y) \frac{1}{\lambda_l - z} dx dy \\ &:= A_1 + A_2 + A_3. \end{aligned} \tag{5.2}$$



Observe that  $|y| \geq 1/2$  in the integral region of the first term  $A_1$ . Because  $(\lambda_l - z)^{-1}$  is analytic in this region, we can use integration by parts. It is not difficult to derive that

$$A_1 = \frac{n}{2\pi} \int_{-2-\epsilon}^{2+\epsilon} \left( \int_{-1}^{-1/2} + \int_{1/2}^1 \right) \left( \sum_{j=0}^3 \frac{1}{j!} i[y^j \sigma'(y)]^{(j)} \right) \varphi(x) s_n(x + iy) dx dy.$$

For the term  $A_2$ , we note that since  $x \in \mathcal{U}_\epsilon \setminus I$ ,  $|\varphi^{(4)}(x)| \leq C$  with some positive constant  $C$  independent of  $n$  by assumption. Moreover, we always have the elementary inequality

$$|(\lambda_l - z)^{-1}| \leq y^{-1}. \quad (5.3)$$

Let  $y_0 = n^{-1-c_0}$ , we decompose the integral region in  $A_2$  into two parts:  $|y| \leq y_0$  and  $|y| > y_0$ . Then (5.3) implies that

$$\left| \frac{1}{2\pi} \sum_{l=1}^n \left( \int_b^{2+\epsilon} + \int_{-2-\epsilon}^a \right) \int_{|y| \leq y_0} \frac{1}{3!} \varphi^{(4)}(x) (iy)^3 \sigma(y) \frac{1}{\lambda_l - z} dx dy \right| \leq C n^{-2-3c_0}.$$

Therefore, we can write

$$A_2 = \frac{n}{2\pi} \left( \int_b^{2+\epsilon} + \int_{-2-\epsilon}^a \right) \int_{|y| > y_0} \frac{1}{3!} \varphi^{(4)}(x) (iy)^3 \sigma(y) s_n(z) dx dy + \mathcal{O}(n^{-2-3c_0}).$$

For the third term  $A_3$ , we will condition on the event  $N_n(I) = 0$ . It is clear that if there is no eigenvalue in the interval  $I = [a, b]$ , then  $(\lambda_l - z)^{-1}$  is continuously differentiable w.r.t  $x$  and  $y$  in the integral region  $I \times [-1, 1]$ . Consequently, when  $N_n(I) = 0$ , we can apply integration by parts to the term  $A_3$  and obtain

$$\begin{aligned} A_3 &= \frac{1}{2\pi} \frac{1}{3!} \sum_{l=1}^n \sum_{\alpha=0}^2 i^{3-\alpha} \int_{-1}^1 (y^3 \sigma(y))^{(\alpha)} \left( \frac{\varphi^{(3-\alpha)}(b)}{\lambda_l - b - iy} - \frac{\varphi^{(3-\alpha)}(a)}{\lambda_l - a - iy} \right) dy \\ &+ \frac{1}{2\pi} \frac{1}{3!} \sum_{l=1}^n \int_a^b \int_{-1}^1 \varphi'(x) (y^3 \sigma(y))^{(3)} \frac{1}{\lambda_l - z} dx dy \\ &= \frac{n}{2\pi} \frac{1}{3!} \sum_{\alpha=0}^2 i^{3-\alpha} \int_{|y| > n^{-4}} (y^3 \sigma(y))^{(\alpha)} \left( \varphi^{(3-\alpha)}(b) s_n(b + iy) - \varphi^{(3-\alpha)}(a) s_n(a + iy) \right) dy \\ &+ \frac{1}{2\pi} \frac{1}{3!} \sum_{l=1}^n \int_a^b \int_{-1}^1 \varphi'(x) (y^3 \sigma(y))^{(3)} \frac{1}{\lambda_l - z} dx dy + \mathcal{O}(n^{-3}). \end{aligned}$$

In the last step, we used (5.3) again to assert

$$\frac{1}{2\pi} \frac{1}{3!} \sum_{l=1}^n \sum_{\alpha=0}^2 i^{3-\alpha} \int_{|y| \leq n^{-4}} (y^3 \sigma(y))^{(\alpha)} \left( \frac{\varphi^{(3-\alpha)}(b)}{\lambda_l - b - iy} - \frac{\varphi^{(3-\alpha)}(a)}{\lambda_l - a - iy} \right) dy = \mathcal{O}(n^{-3}).$$

Consequently, when  $N_n(I) = 0$ , we obtain

$$\begin{aligned}
\mathcal{L}_n[\varphi] &= \frac{n}{2\pi} \int_{-2-\epsilon}^{2+\epsilon} \left( \int_{-1}^{-1/2} + \int_{1/2}^1 \right) \left( \sum_{j=0}^3 \frac{1}{j!} i(y^j \sigma'(y))^{(j)} \right) \varphi(x) s_n(x+iy) dx dy \\
&+ \frac{n}{2\pi} \left( \int_b^{2+\epsilon} + \int_{-2-\epsilon}^a \right) \int_{|y|>y_0} \frac{1}{3!} \varphi^{(4)}(x) (iy)^3 \sigma(y) s_n(x+iy) dx dy \\
&+ \frac{n}{2\pi} \frac{1}{3!} \sum_{\alpha=0}^2 i^{3-\alpha} \int_{|y|>n^{-4}} (y^3 \sigma(y))^{(\alpha)} \left( \varphi^{(3-\alpha)}(b) s_n(b+iy) - \varphi^{(3-\alpha)}(a) s_n(a+iy) \right) dy \\
&+ \frac{n}{2\pi} \frac{1}{3!} \int_a^b \int_{|y|>n^{-5}} \varphi'(x) (y^3 \sigma(y))^{(3)} s_n(x+iy) dx dy \\
&+ \frac{1}{2\pi} \frac{1}{3!} \sum_{l=1}^n \int_a^b \int_{|y|\leq n^{-5}} \varphi'(x) (y^3 \sigma(y))^{(3)} \frac{1}{\lambda_l - z} dx dy + \mathcal{O}(n^{-2-3c_0}) \\
&:= A_4 + A_5 + A_6 + A_7 + A_8 + \mathcal{O}(n^{-2-3c_0}). \tag{5.4}
\end{aligned}$$

To use the above representation in the proof of Theorem 5.1, we shall provide a more easily handled condition on the Stieltjes transform instead of  $N_n(I) = 0$ . Such a trick is from Tao and Vu [32].

**Lemma 5.4.** *For some positive constant  $A_0$  (independent of  $c_0$  and  $a$ ), if*

$$Im s_n(a + in^{-1-2A_0c_0}) \leq n^{-A_0c_0}, \tag{5.5}$$

*one has*

$$\inf_{x \in I} \min_l |\lambda_l - x| > n^{-1-A_0c_0} \tag{5.6}$$

*when  $c_0$  is sufficiently small.*

A direct consequence of Lemma 5.4 is  $N_n(I) = 0$  when (5.5) holds and  $c_1 > A_0c_0$ .

*Proof.* By definition, one has

$$Im s_n(a + in^{-1-2A_0c_0}) = n^{-2-2A_0c_0} \sum_{l=1}^n \frac{1}{(\lambda_l - a)^2 + n^{-2-4A_0c_0}}.$$

By the assumption  $Im s_n(a + in^{-1-2A_0c_0}) \leq n^{-A_0c_0}$ , we can get

$$\min_l (\lambda_l - a)^2 \geq n^{-2-A_0c_0} - n^{-2-4A_0c_0}.$$

Consequently, when  $n$  is sufficiently large one has

$$\min_l |\lambda_l - a| \geq 2n^{-1-A_0c_0}. \tag{5.7}$$

Now if  $c_0$  is sufficiently small such that  $c_1 \geq A_0c_0$ , we can easily get by triangular inequality that

$$\min_{x \in I} \min_l |\lambda_l - x| \geq n^{-1-A_0c_0}.$$

Here  $c_1$  is the constant in (b) of the definition of  $\mathcal{F}_n$ . Thus we conclude the proof.  $\square$

Moreover, we have the following lemma due to Tao and Vu [32].

**Lemma 5.5.** (Corollary 15, [32]) *For any  $a \in [-2+\delta, 2-\delta]$ , there exists a sufficiently large constant  $A_0 > 0$  (independent of  $c_0$  and  $a$ ),*

$$\text{Im}s_n(a + in^{-1-2A_0c_0}) \leq n^{-A_0c_0}/2$$

*holds with high probability.*

**Remark 5.6.** *The proof of the above lemma in [32] is based on the level repulsion estimate of Wigner matrices (see Proposition 14, [32]). The proof of the level repulsion in [32] needs the conditions that  $a \in [-2+\delta, 2-\delta]$  and the distributions of the matrix elements are supported on at least three points. That is why we make these assumptions in our main results. However, we believe these restrictions are not necessary and can be removed. We will not pursue this direction in this paper.*

Pursuing the argument in [32], we define a smooth cutoff function  $\chi(x)$  to the region  $|x| \leq n^{-A_0c_0}$  that equals 1 for  $|x| \leq n^{-A_0c_0}/2$ . Thus by Lemma 5.5, one sees that  $\chi(\text{Im}s_n(a + in^{-1-2A_0c_0}))$  is equal to 1 with high probability. Consequently, it suffices to prove the fact that the quantity

$$\mathbb{E}\{G(\mathcal{L}_n[\varphi])\chi(\text{Im}s_n(a + in^{-1-2A_0c_0}))\}$$

is highly insensitive for  $\varphi \in \mathcal{F}_n$ .

Moreover, we have mentioned above that by Lemma 5.4, one has  $\chi(\text{Im}s_n(a + in^{-1-2A_0c_0})) \neq 0$  implies that  $N_n(I) = 0$ . Therefore, we can use the representation (5.4) for  $\mathcal{L}_n[\varphi]$ . Moreover, by the bound on the derivative of  $G$  (see (5.1)), one has

$$\begin{aligned} & \mathbb{E}\{G(\mathcal{L}_n[\varphi])\chi(\text{Im}s_n(a + in^{-1-2A_0c_0}))\} \\ &= \mathbb{E}\{G(\sum_{l=4}^8 A_l)\chi(\text{Im}s_n(a + in^{-1-2A_0c_0}))\} + \mathcal{O}(n^{-2-2c_0}). \end{aligned} \quad (5.8)$$

To show that the above quantity is highly insensitive, the main task is to provide the stability of  $s_n(x+iy)$  involved in  $A_i$  and  $\chi(\text{Im}s_n(a + in^{-1-2A_0c_0}))$  in the swapping procedure. To this end, we need the Taylor expansion for  $s_n(x+iy)$  proved by Tao and Vu [32]: Lemma 2.1 stated in Section 2. That is to say, for Wigner matrices  $M_n$  and  $M'_n$ , we can start from  $M_n$ , and then replace its elements one pair (or one unit for the diagonal case) a time by the corresponding one of  $M'_n$  and study the stability of  $s_n(x+iy)$  under such a swapping process. To achieve this aim, we let  $M_n^{(1)}, M_n^{(2)}$  be two adjacent matrices in the swapping procedure in the sense that we can write

$$M^{(1)} = M_0 + \frac{1}{\sqrt{n}}\xi^{(1)}V, \quad M^{(2)} = M_0 + \frac{1}{\sqrt{n}}\xi^{(2)}V$$

for some elementary matrix  $V$ . And  $\xi^{(1)}, \xi^{(2)}$  are two real random variables matching to the fourth order and bounded in magnitude by  $n^{\mathcal{O}(c_0)}$ . Moreover,  $M_0$  is independent of  $\xi^{(1)}$  and  $\xi^{(2)}$ .

To describe the swapping process, we use the notation  $s_{\xi^{(1)}}(x+iy)$  to denote the Stieltjes transform for  $M_n^{(1)}$ , and  $s_{\xi^{(2)}}(x+iy)$  for  $M_n^{(2)}$ . Correspondingly we distinguish  $\mathcal{L}_n[\varphi]$  by  $\mathcal{L}_n^{(1)}[\varphi]$  and  $\mathcal{L}_n^{(2)}[\varphi]$  for  $M_n^{(1)}$  and  $M_n^{(2)}$  respectively. Besides,

we use the notation  $A_l^{(1)}$  and  $A_l^{(2)}$  to denote  $A_l$  ( $l = 4, \dots, 8$ ) for  $M_n^{(1)}$  and  $M_n^{(2)}$  respectively. With these notations, we will show the quantity

$$\mathbb{E}\{G(\sum_{l=4}^8 A_l^{(1)})\chi(Im s_{\xi^{(1)}}(a + in^{-1-2A_0c_0}))\}$$

only changes by  $\mathcal{O}(n^{-2-\mathcal{O}(c_0)})$  when  $\xi^{(1)}$  is replaced by  $\xi^{(2)}$  in the off diagonal case, or  $\mathcal{O}(n^{-1-\mathcal{O}(c_0)})$  in the diagonal case. Then by a telescoping arguments, after  $\mathcal{O}(n^2)$  steps of replacement, we can easily get that

$$\mathbb{E}\{G(\mathcal{L}_n[\varphi])\chi(Im s_n(a + in^{-1-2A_0c_0}))\}$$

is highly insensitive.

In order to apply Lemma 2.1, one shall guarantee the condition (2.9). We need the following crucial lemma.

**Lemma 5.7.** (*Uniform resolvent bounds*). *We have the following two assertions on the resolvent bounds.*

(i): *If  $\chi(Im s_{\xi^{(1)}}(a + in^{-1-2A_0c_0})) \neq 0$ , then with overwhelming probability*

$$\sup_{x \in I} \sup_{y > 0} \|R_{\xi^{(1)}}(x + iy)\|_{(\infty,1)} = \mathcal{O}(n^{\mathcal{O}(c_0)}) \quad (5.9)$$

and

$$\sup_{x \in I} \sup_{y > 0} \|R_0(x + iy)\|_{(\infty,1)} = \mathcal{O}(n^{\mathcal{O}(c_0)}). \quad (5.10)$$

(ii): *If  $y_0 = n^{-1-c_0}$ , then with overwhelming probability*

$$\sup_{x \in \mathcal{U}_\epsilon} \sup_{|y| > y_0} \|R_{\xi^{(1)}}(x + iy)\|_{(\infty,1)} = \mathcal{O}(n^{\mathcal{O}(c_0)}) \quad (5.11)$$

and

$$\sup_{x \in \mathcal{U}_\epsilon} \sup_{|y| > y_0} \|R_0(x + iy)\|_{(\infty,1)} = \mathcal{O}(n^{\mathcal{O}(c_0)}). \quad (5.12)$$

*Proof.* At first, we prove (i). We learn from the proof of Lemma 5.4 that when

$$\chi(Im s_{\xi^{(1)}}(a + in^{-1-2A_0c_0})) \neq 0,$$

there exists (5.7). Besides, by the spectral decomposition we can easily get

$$\|R_{\xi^{(1)}}(x + iy)\|_{(\infty,1)} \leq \sum_{l=1}^n \frac{\sup_{1 \leq j \leq n} \|u_j(M_n^{(1)})\|_{l^\infty}^2}{|\lambda_l(M_n^{(1)}) - x - iy|}, \quad (5.13)$$

where  $u_j(M_n^{(1)})$  is the unit eigenvector of  $M_n^{(1)}$  corresponding to  $\lambda_j(M_n^{(1)})$ . With (7.2) in Lemma 7.1, one has

$$\|R_{\xi^{(1)}}(x + iy)\|_{(\infty,1)} \leq n^{-1+\mathcal{O}(c_0)} \sum_{l=1}^n \frac{1}{|\lambda_l(M_n^{(1)}) - x|}.$$

For  $x \in I$ , by using (5.7) we have

$$|\lambda_l(M_n^{(1)}) - x| \geq |\lambda_l(M_n^{(1)}) - a| - n^{-1-c_1} \geq \frac{1}{2}|\lambda_l(M_n^{(1)}) - a|$$

when  $c_1 > A_0 c_0$ . Consequently we have

$$\sup_{x \in I} \sup_{y > 0} \|R_{\xi^{(1)}}(x + iy)\|_{(\infty, 1)} \leq 2n^{-1+\mathcal{O}(c_0)} \sum_{l=1}^n \frac{1}{|\lambda_l(M_n^{(1)}) - a|}.$$

By the argument in Tao and Vu (the proof of Lemma 16, [32]), if  $\chi(\text{Im}s_{\xi^{(1)}}(a + in^{-1-2A_0 c_0})) \neq 0$ , one has with overwhelming probability that

$$\sum_{l=1}^n \frac{1}{|\lambda_l(M_n^{(1)}) - a|} = \mathcal{O}(n^{1+\mathcal{O}(c_0)}).$$

Thus we have

$$\sup_{x \in I} \sup_{y > 0} \|R_{\xi^{(1)}}(x + iy)\|_{(\infty, 1)} = \mathcal{O}(n^{\mathcal{O}(c_0)}).$$

For  $R_0(x + iy)$ , we use the fact that for  $y > 0$ , when  $|t| \|R_t\|_{(\infty, 1)} = o(\sqrt{n})$ ,

$$\|R_0\|_{(\infty, 1)} \leq (1 + o(1)) \|R_t\|_{(\infty, 1)}. \quad (5.14)$$

(5.14) is a consequence of Neumann series formula, we refer to [32] for the details of the proof. Thus we also have

$$\sup_{x \in I} \sup_{y > 0} \|R_0(x + iy)\|_{(\infty, 1)} = \mathcal{O}(n^{\mathcal{O}(c_0)}).$$

Now we turn to the proof of (ii). For a sufficiently large constant  $A$ , we set  $\eta = n^{-1+A c_0}$ . We cover the interval  $\mathcal{U}_\epsilon$  by the union of the intervals  $I_k = [(k - \frac{1}{2})\eta, (k + \frac{1}{2})\eta]$  with the integer index  $k$  running from  $-\lfloor(2 + \epsilon)\eta^{-1}\rfloor - 1$  to  $\lfloor(2 + \epsilon)\eta^{-1}\rfloor + 1$ .

Now note that by (5.13) we have

$$\begin{aligned} \|R_{\xi^{(1)}}(x + iy)\|_{(\infty, 1)} &\leq n^{-1+\mathcal{O}(c_0)} \sum_{l=1}^n \frac{1}{\max\{|\lambda_l(M_n^{(1)}) - x|, |y|\}} \\ &= n^{-1+\mathcal{O}(c_0)} \sum_k \sum_{l: \lambda_l(M_n^{(1)}) \in I_k} \frac{1}{\max\{|\lambda_l(M_n^{(1)}) - x|, |y|\}}. \end{aligned}$$

By the fact that  $N_J = O(n|J|)$  with overwhelming probability for any interval  $J$  with length  $|J| \geq n^{-1+A c_0}$  (see Lemma 7.1), we can immediately get

$$\sup_{x \in \mathcal{U}_\epsilon} \sup_{y > y_0} \|R_{\xi^{(1)}}(x + iy)\|_{(\infty, 1)} = \mathcal{O}(n^{\mathcal{O}(c_0)})$$

with overwhelming probability. Again by the resolvent bound (5.14) one has

$$\sup_{x \in \mathcal{U}_\epsilon} \sup_{y > y_0} \|R_0(x + iy)\|_{(\infty, 1)} = \mathcal{O}(n^{\mathcal{O}(c_0)}).$$

Thus we complete the proof.  $\square$

If we condition on the event that (5.10) and (5.12) hold, then we also have (5.9) and (5.11) by swapping the roles of  $R_0$  and  $R_t$  in (5.14). Since  $|\xi^{(1)}| \leq n^{\mathcal{O}(c_0)}$ ,

together with (5.10) and (5.12), we can use Lemma 2.1. Besides, under the condition  $\chi(Im s_{\xi^{(1)}}(a + in^{-1-2A_0c_0})) \neq 0$ , we have (5.6), which trivially implies

$$\sup_{x \in I} \sup_y \left| \frac{1}{n} \sum_{l=1}^n \frac{1}{\lambda_l - z} \right| = \mathcal{O}(n^{1+\mathcal{O}(c_0)}).$$

Consequently, we have

$$A_8^{(1)} = \mathcal{O}(n^{-3+\mathcal{O}(c_0)}).$$

Now we set

$$\begin{aligned} A_4^0 &= \frac{n}{2\pi} \int_{-2-\epsilon}^{2+\epsilon} \left( \int_{-1}^{-1/2} + \int_{1/2}^1 \right) \left( \sum_{j=0}^3 \frac{(-1)^j}{j!} i(y^j \sigma'(y))^{(j)} \right) \varphi(x) s_0(x + iy) dx dy, \\ A_5^0 &= \frac{n}{2\pi} \left( \int_b^{2+\epsilon} + \int_{-2-\epsilon}^a \right) \int_{|y| > y_0} \frac{1}{3!} \varphi^{(4)}(x) (iy)^3 \sigma(y) s_0(x + iy) dx dy, \\ A_6^0 &= \frac{n}{2\pi} \frac{1}{3!} \sum_{\alpha=0}^2 i^{3-\alpha} \int_{|y| > n^{-4}} (y^3 \sigma(y))^{(\alpha)} \left( \varphi^{(3-\alpha)}(b) s_0(b + iy) - \varphi^{(3-\alpha)}(a) s_0(a + iy) \right) dy, \\ A_7^0 &= \frac{n}{2\pi} \frac{1}{3!} \int_a^b \int_{|y| > n^{-5}} \varphi'(x) (y^3 \sigma(y))^{(3)} s_0(x + iy) dx dy. \end{aligned}$$

Then we let

$$\begin{aligned} G_n^{(1)} &=: G \left( \sum_{l=4}^7 A_l^{(1)} \right) \chi(Im s_{\xi^{(1)}}(a + in^{-1-2A_0c_0})), \\ G_n^0 &=: G \left( \sum_{l=4}^7 A_l^0 \right) \chi(Im s_0(a + in^{-1-2A_0c_0})). \end{aligned}$$

Correspondingly we can define  $G_n^{(2)}$ . Our aim is to expand  $G_n^{(1)}$  around  $G_n^{(0)}$ . We formulate the result as the following lemma.

**Lemma 5.8.** *With the above notations, when we condition on the event that (5.10) and (5.12) hold, we have*

$$G_n^{(1)} = G_n^0 + \sum_{j=1}^4 d_j (\xi^{(1)})^j + \mathcal{O}(n^{-5/2+\mathcal{O}(c_0)}), \quad (5.15)$$

where the coefficients  $d_j$  are independent of  $\xi^{(1)}$  and obey the bounds

$$d_j = \mathcal{O}(n^{-j/2+\mathcal{O}(c_0)}), \quad j = 1, \dots, 4.$$

*Proof.* Note that if (5.10) and (5.12) hold, we can use Lemma 2.1 to expand  $s_{\xi^{(1)}}$  (resp.  $A_l^{(1)}$ ) around  $s_0$  (resp.  $A_l^0$ ). Then by the assumption on the derivatives of  $G(x)$  and the fact that

$$\frac{d^j}{dx^j} \chi(x) = \mathcal{O}(n^{\mathcal{O}(jc_0)}),$$

we can conclude (5.15) by applying Taylor expansion to  $G(\cdot)$  and  $\chi(\cdot)$ .  $\square$

*Proof of Theorem 5.1.* Using Lemma 5.8 to both  $M^{(1)}$  and  $M^{(2)}$ , we have

$$G_n^{(1)} = G_n^0 + \sum_{j=1}^4 d_j(\xi^{(1)})^j + \mathcal{O}(n^{-5/2+\mathcal{O}(c_0)})$$

and

$$G_n^{(2)} = G_n^0 + \sum_{j=1}^4 d_j(\xi^{(2)})^j + \mathcal{O}(n^{-5/2+\mathcal{O}(c_0)}).$$

Then taking expectation with respect to  $\xi^{(1)}$  and  $\xi^{(2)}$  respectively, and by the telescope arguments on  $\mathcal{O}(n^2)$  steps of swapping, we can immediately get the conclusion by the matching moments assumption.  $\square$

As an application of Theorem 5.1, we can prove our main result Theorem 1.7.

*Proof of Theorem 1.7.* Note that Lemma 3.2 holds for Wigner matrices under the conditions of Theorem 1.7. By the argument in Section 3, it suffices to show that (ii) and (iii) of Theorem 1.1 are still valid for general Wigner matrices. To combine Theorem 1.1 with Theorem 5.1, we let  $M_n$  be a general Wigner matrix and  $M'_n$  be GUE. Moreover, we shall truncate the elements of  $W_n$  and  $W'_n$  at  $\mathcal{O}(n^{c_0})$  to adapt to the condition of Theorem 5.1. Under the condition  $\mathbf{C}_0$ , it is easy to see such a truncation does not alter the limiting behavior of both two types of partial linear eigenvalue statistics. Moreover, the truncation will change the first four moments of the elements by only  $\mathcal{O}(e^{-n^{\mathcal{O}(c_0)}})$ , which can be absorbed in the remainder  $\mathcal{O}(n^{-5/2+\mathcal{O}(c_0)})$  when we take expectations on both sides of (5.15). Next, we define a smooth modification of  $f_u$  by

$$h_u(x) =: h_u(n, x) = (f(x) - f(u))\tilde{\chi}_u(n, x), \quad (5.16)$$

where  $\tilde{\chi}_u(n, x)$  is an  $n$ -dependent smooth cutoff to the region  $x \in (-\infty, u+n^{-1-c_1}/2)$  that equals 1 for  $x \in (-\infty, u-n^{-1-c_1}/2)$ . Similar to (3.3), one has

$$\mathcal{A}_n[f_u; u] = \mathcal{L}_n[f_u] = \mathcal{L}_n[h_u] + o(1)$$

with overwhelming probability for all Wigner matrices satisfying *Condition  $\mathbf{C}_0$* .

It is not difficult to see that to extend (ii) in Theorem 1.1, it suffices to show that for any interval  $J = [j_1, j_2]$ ,

$$\mathbb{P}(\mathcal{L}_n^{M'}[h_u] \in J_-) - n^{-c_0} \leq \mathbb{P}(\mathcal{L}_n^M[h_u] \in J) \leq \mathbb{P}(\mathcal{L}_n^{M'}[h_u] \in J_+) + n^{-c_0}. \quad (5.17)$$

Here

$$J_+ = [j_1 - n^{-c_0/10}, j_2 + n^{-c_0/10}], \quad J_- = [j_1 + n^{-c_0/10}, j_2 - n^{-c_0/10}].$$

We only show the second inequality of (5.17) below since the first one is analogous. Let  $G: \mathbb{R} \rightarrow \mathbb{R}^+$  equal to one in  $J$  and vanish outside of  $J_+$  such that

$$\frac{d^j}{dx^j} G(x) = \mathcal{O}(n^{c_0}), \quad j = 0, \dots, 5.$$

We can apply Theorem 5.1 to  $G$  defined above. Observe that

$$\mathbb{P}(\mathcal{L}_n^M[h_u] \in J) \leq \mathbb{E}G(\mathcal{L}_n^M[h_u]),$$

and

$$\mathbb{E}G(\mathcal{L}_n^{M'}[h_u]) \leq \mathbb{P}(\mathcal{L}_n^{M'}[h_u] \in J_+).$$

Thus by Theorem 5.1, we immediately get the second inequality of (5.17). Thus we have proved that (ii) of Theorem 1.1 is still valid for general Wigner matrices. For (iii), by using (3.23) and (3.24) again, we can also reduce the problem to compare  $\mathcal{L}_n^M[f_{\gamma_{k/n}}]$  and  $\mathcal{L}_n^{M'}[f_{\gamma_{k/n}}]$ . Then we only need to replace  $u$  by  $\gamma_{k/n}$  in the arguments above and get the conclusion. Thus we finally complete the proof of Theorem 1.7.  $\square$

## 6. PARTIAL SUM PROCESS FOR WIGNER MATRICES

In this section, we prove Theorem 1.8 by providing the finite-dimensional convergence and tightness of the sequence  $\mathcal{S}_n[f; t]$ . Our strategy is to use the comparison theorem established in the last section to both two parts of the proof. Below we will use the notation  $\mathcal{S}_n^M[f; t]$  and  $\mathcal{S}_n^{M'}[f; t]$  to denote  $\mathcal{S}_n[f; t]$  for Wigner matrices  $M_n$  and  $M'_n$  respectively. Moreover, we will specify  $M'_n$  to be GUE in this section. At first, we will prove the following lemma on the expectations.

**Lemma 6.1.** *Under the assumption of Theorem 1.8, for  $t \in [\delta, 1 - \delta]$ , we have*

$$\mathbb{E}\mathcal{S}_n^M[f; t] - \mathbb{E}\mathcal{S}_n^{M'}[f; t] = \mathcal{O}(n^{-c})$$

for some fixed constant  $c > 0$ .

*Proof.* By the assumptions on  $f(x)$  in Theorem 1.8 and the large deviation estimate of extreme eigenvalue in Lemma 7.4, we can and do assume that  $f(x)$  is compactly supported on  $\mathcal{U}_\epsilon$  below. By definition, we shall provide that

$$\mathbb{E}\mathcal{B}_n^M[f; \lfloor nt \rfloor] - \mathbb{E}\mathcal{B}_n^{M'}[f; \lfloor nt \rfloor] = \mathcal{O}(n^{-c}), \quad (6.1)$$

and

$$\mathbb{E}f(\lambda_{\lfloor nt \rfloor + 1}(M_n)) - \mathbb{E}f(\lambda_{\lfloor nt \rfloor + 1}(M'_n)) = \mathcal{O}(n^{-c}). \quad (6.2)$$

Relying on (3.23) and the fact that both  $\mathcal{B}_n^M[f; \lfloor nt \rfloor]$  and  $\mathcal{B}_n^{M'}[f; \lfloor nt \rfloor]$  are  $\mathcal{O}(n)$ , one can prove

$$\mathbb{E}\widehat{\mathcal{B}}_n^M[f; \lfloor nt \rfloor] - \mathbb{E}\widehat{\mathcal{B}}_n^{M'}[f; \lfloor nt \rfloor] = \mathcal{O}(n^{-c})$$

instead of (6.1). Moreover, by (3.24) and the definition of  $h_u(x)$  in (5.16), apparently it suffices to show

$$\mathbb{E}\mathcal{L}_n^M[h_{u(t)}] - \mathbb{E}\mathcal{L}_n^{M'}[h_{u(t)}] = \mathcal{O}(n^{-c}). \quad (6.3)$$

Here we recall the notation  $u(t) = \gamma_{\lfloor nt \rfloor/n}$ .

Let  $\chi_3(x)$  be a smooth cutoff function to the region  $[-2n^{c_0/2}, 2n^{c_0/2}]$  which equals to 1 in  $[-n^{c_0/2}, n^{c_0/2}]$ . Now we define the function  $G_1$  as

$$G_1(x) = (x - \sum_{l=1}^n h_{u(t)}(\gamma_{l/n}))\chi_3(x - \sum_{l=1}^n h_{u(t)}(\gamma_{l/n})).$$



Obviously,  $G_1(x)$  satisfies the condition (5.1).

Now we claim that

$$\mathbb{E}G_1(\mathcal{L}_n[h_{u(t)}]) = \mathbb{E}\mathcal{L}_n[h_{u(t)}] - \sum_{l=1}^n h_{u(t)}(\gamma_{l/n}) + \mathcal{O}(n^{-c}). \quad (6.4)$$

To see (6.4), it suffices to show that

$$\chi_3(\mathcal{L}_n[h_{u(t)}] - \sum_{l=1}^n h_{u(t)}(\gamma_{l/n})) = 1 \quad (6.5)$$

holding with overwhelming probability. To this end, we use the rigidity property stated in Lemma 7.2. With the aid of this lemma, we now show the validity of (6.5) as follows. Note that with overwhelming probability one has

$$\begin{aligned} |\mathcal{L}_n[h_{u(t)}] - \sum_{l=1}^n h_{u(t)}(\gamma_{l/n})| &\leq \sup_{x \in \mathcal{U}_\epsilon} |h'_{u(t)}(x)| \sum_{l=1}^n |\lambda_l - \gamma_{l/n}| \\ &\leq C(\log n)^{C \log \log n} \sum_{l=1}^n [\min(l, n-l+1)]^{-1/3} n^{-2/3} \leq n^{c_0/2} \end{aligned}$$

for  $n$  sufficiently large. Thus we have shown (6.4).

Consequently, it suffices to show (6.2) and

$$\mathbb{E}G_1(\mathcal{L}_n^M[h_{u(t)}]) - \mathbb{E}G_1(\mathcal{L}_n^{M'}[h_{u(t)}]) = \mathcal{O}(n^{-c}). \quad (6.6)$$

Observe that (6.6) is a direct consequence of Theorem 5.1 (After a harmless truncation towards the elements of both  $W_n$  and  $W'_n$ ). Thus we only need to show (6.2) in the sequel. Note that

$$\begin{aligned} &|f(\lambda_{\lfloor nt \rfloor + 1}(M_n)) - f(\lambda_{\lfloor nt \rfloor + 1}(M'_n))| \\ &\leq \sup_{x \in \mathcal{U}_\epsilon} |f'(x)| |\lambda_{\lfloor nt \rfloor + 1}(M_n) - \lambda_{\lfloor nt \rfloor + 1}(M'_n)| \\ &\leq C \frac{(\log n)^{C \log \log n}}{n} \end{aligned}$$

holding with overwhelming probability. By the assumption that  $f(x)$  is compactly supported thus bounded, one can immediately get (6.2). Hence, we complete the proof.  $\square$

Now we begin to prove the finite dimensional convergence of the sequence  $(\mathcal{S}_n^M[f; t])^\circ$ . We formulate the result as the following lemma and then prove it.

**Lemma 6.2.** *Under the assumption of Theorem 1.8, for any fixed positive integer  $r$  and points  $t_1, \dots, t_r \in [\delta, 1 - \delta]$ , and for any fixed numbers  $\alpha_1, \dots, \alpha_r \in \mathbb{R}$ , we have*

$$\sum_{l=1}^r \alpha_l (\mathcal{S}_n^M[f; t_l])^\circ \xrightarrow{d} \sum_{l=1}^r \alpha_l \mathcal{S}[f; t_l].$$

*Proof.* By the same discussion as that for GUE taken in Section 4, we can transfer the problem to show

$$(\mathcal{L}_n^M[\sum_{l=1}^r \alpha_l f_{\gamma_{t_l}}])^\circ \xrightarrow{d} \sum_{l=1}^r \alpha_l \mathcal{S}[f; t_l]$$

instead. As mentioned in Section 4,  $\sum_{l=1}^r \alpha_l f_{\gamma_{t_l}}$  is a continuous function with  $r$  possibly non-differentiable points  $t_1, \dots, t_r$ . Now we choose  $r$  intervals  $J_1, \dots, J_r$  containing  $\gamma_{t_1}, \dots, \gamma_{t_r}$  respectively with lengths  $|J_l| \leq n^{-1-c_1}$  for all  $l = 1, \dots, r$ . And then we define a smooth modification function  $h_{t_1, \dots, t_r}$  which coincides with  $\sum_{l=1}^r \alpha_l f_{\gamma_{t_l}}$  on  $\mathbb{R} \setminus \cup_{l=1}^r J_l$ . Thus, it is easy to see  $h_{t_1, \dots, t_r} \in \mathcal{F}_n^r$ . Now by using Theorem 5.1 and a routine discussion as that in the proof of Theorem 1.7, one can get that

$$\mathcal{L}_n^M[h_{t_1, \dots, t_r}] - \mathbb{E}\mathcal{L}_n^{M'}[h_{t_1, \dots, t_r}] \xrightarrow{d} \sum_{l=1}^r \alpha_l \mathcal{S}[f; t_l]. \quad (6.7)$$

Here  $M' = M'_n$  is GUE. Apparently, by (6.7) one can get

$$\mathcal{L}_n^M[\sum_{l=1}^r \alpha_l f_{\gamma_{t_l}}] - \mathbb{E}\mathcal{L}_n^{M'}[\sum_{l=1}^r \alpha_l f_{\gamma_{t_l}}] \xrightarrow{d} \sum_{l=1}^r \alpha_l \mathcal{S}[f; t_l].$$

Then by Lemma 6.1, we can conclude the proof of Lemma 6.2.  $\square$

Thus the only thing left is to prove the tightness of the sequence  $(\mathcal{S}_n^M[f; t])^\circ$ . It suffices to provide the Hölder condition: for arbitrary  $t, s \in [\delta, 1 - \delta]$  and sufficiently large  $n$

$$\mathbb{E}|(\mathcal{S}_n^M[f; t])^\circ - (\mathcal{S}_n^M[f; s])^\circ|^2 \leq C|t - s|^\alpha$$

holds for some constants  $C > 0$ , and  $\alpha > 1$  both independent of  $t, s$ .

At first, we claim that under the assumptions in Theorem 1.8, we always have for  $l \in [\delta n, (1 - \delta)n]$

$$\text{Var}\{f(\lambda_l)\} \leq C \frac{\log n}{n^2} \quad (6.8)$$

with some positive constant  $C$  depending only on  $\delta$  and  $f$ . Based on Lemma 7.5, the proof of (6.8) is analogous to that of (4.1). The only difference is that we need to use Lemma 7.4 instead of (2.6). Here we omit the detail.

Similar to the GUE case, we always assume that  $s \leq t$  and separate the issue into three cases:  $t - s \leq n^{-1}$ ,  $n^{-1} \leq t - s \leq \eta_n$  and  $t - s \geq \eta_n$ . It is easy to see the discussions for the first two cases for GUE in Section 4 are also valid for general Wigner matrix with the aid of (6.8). Thus we will focus on the third case in the sequel.

**Lemma 6.3.** *Under the assumptions of Theorem 1.8, for  $t, s \in [\delta, 1 - \delta]$  such that  $t - s \geq \eta_n =: \log^{-L} n$  with sufficiently large constant  $L$ , one has for  $n$  large enough*

$$\mathbb{E}|(\mathcal{S}_n^M[f; t])^\circ - (\mathcal{S}_n^M[f; s])^\circ|^2 \leq C|t - s|^\alpha$$

for some constants  $C > 0$ , and  $\alpha > 1$  both independent of  $t, s$ .

*Proof.* To prove Lemma 6.3, we will use Theorem 5.1 to compare the general case with the Gaussian case. At first, note that it suffices to prove

$$\mathbb{E}|(\mathcal{B}_n^M[f; \lfloor nt \rfloor])^\circ - (\mathcal{B}_n^M[f; \lfloor ns \rfloor])^\circ|^2 \leq C|t - s|^\alpha$$

instead since

$$\text{Var}f(\lambda_{\lfloor nt \rfloor + 1}) + \text{Var}f(\lambda_{\lfloor ns \rfloor + 1}) \leq C(\delta) \frac{\log^2 n}{n}.$$

Moreover, by (2.18) we have

$$\begin{aligned} & \mathbb{E}|(\mathcal{B}_n^M[f; \lfloor nt \rfloor])^\circ - (\mathcal{B}_n^M[f; \lfloor ns \rfloor])^\circ|^2 \\ &= \mathbb{E}|(\widehat{\mathcal{B}}_n^M[f; \lfloor nt \rfloor])^\circ - (\widehat{\mathcal{B}}_n^M[f; \lfloor ns \rfloor])^\circ|^2 + \mathcal{O}\left(\frac{(\log n)^{C \log \log n}}{n}\right) \\ &= \text{Var}\{\mathcal{L}_n^M[f_{u(t)} - f_{u(s)}]\} + \mathcal{O}\left(\frac{(\log n)^{C \log \log n}}{n}\right) \\ &= \text{Var}\{\mathcal{L}_n^M[h_{u(t)} - h_{u(s)}]\} + \mathcal{O}(n^{-c}). \end{aligned} \quad (6.9)$$

Here  $h_u$  is defined in (5.16).

Note that we have (4.3) for GUE. Thus by (6.9), we see that it suffices to show

$$\text{Var}\{\mathcal{L}_n^M[h_{u(t)} - h_{u(s)}]\} - \text{Var}\{\mathcal{L}_n^{M'}[h_{u(t)} - h_{u(s)}]\} = \mathcal{O}(n^{-c}). \quad (6.10)$$

Observe that

$$\begin{aligned} \text{Var}\{\mathcal{L}_n^M[h_{u(t)} - h_{u(s)}]\} &= \mathbb{E}(\mathcal{L}_n^M[h_{u(t)} - h_{u(s)}] - \mathbb{E}\mathcal{L}_n^{M'}[h_{u(t)} - h_{u(s)}])^2 \\ &\quad - (\mathbb{E}\{\mathcal{L}_n^M[h_{u(t)} - h_{u(s)}]\} - \mathbb{E}\{\mathcal{L}_n^{M'}[h_{u(t)} - h_{u(s)}]\})^2 \\ &= \mathbb{E}(\mathcal{L}_n^M[h_{u(t)} - h_{u(s)}] - \mathbb{E}\mathcal{L}_n^{M'}[h_{u(t)} - h_{u(s)}])^2 + \mathcal{O}(n^{-2c}). \end{aligned}$$

Here the last step follows from (6.6). Thus it suffices to show that

$$\mathbb{E}(\mathcal{L}_n^M[h_{u(t)} - h_{u(s)}] - \mathbb{E}\mathcal{L}_n^{M'}[h_{u(t)} - h_{u(s)}])^2 - \text{Var}\{\mathcal{L}_n^{M'}[h_{u(t)} - h_{u(s)}]\} = \mathcal{O}(n^{-2c}).$$

Now we set

$$G_2(x) = (x - \mathbb{E}\mathcal{L}_n^{M'}[h_{u(t)} - h_{u(s)}])\chi_3(x - \mathbb{E}\mathcal{L}_n^{M'}[h_{u(t)} - h_{u(s)}]).$$

Note that again by the rigidity property stated in Lemma 7.2, one has

$$|\mathcal{L}_n^M[h_{u(t)} - h_{u(s)}] - \sum_{l=1}^n (h_{u(t)} - h_{u(s)})(\gamma_{l/n})| \leq (\log n)^{C \log \log n}$$

with overwhelming probability. Simultaneously, one can get that

$$|\mathbb{E}\mathcal{L}_n^{M'}[h_{u(t)} - h_{u(s)}] - \sum_{l=1}^n (h_{u(t)} - h_{u(s)})(\gamma_{l/n})| \leq (\log n)^{C \log \log n}.$$

Thus

$$\chi_3(\mathcal{L}_n^M[h_{u(t)} - h_{u(s)}] - \mathbb{E}\mathcal{L}_n^{M'}[h_{u(t)} - h_{u(s)}]) = 1$$

with overwhelming probability. Consequently, it suffices to compare

$$\mathbb{E}\{G_2(\mathcal{L}_n^M[h_{u(t)} - h_{u(s)}])\}$$

and

$$\mathbb{E}\{G_2(\mathcal{L}_n^{M'}[h_{u(t)} - h_{u(s)}])\}.$$

By using Theorem 5.1 again, we can get

$$\mathbb{E}\{G_2(\mathcal{L}_n^M[h_{u(t)} - h_{u(s)}])\} - \mathbb{E}\{G_2(\mathcal{L}_n^{M'}[h_{u(t)} - h_{u(s)}])\} = \mathcal{O}(n^{-c}).$$

Thus (6.10) follows. Moreover, the remainder  $\mathcal{O}(n^{-c})$  in (6.10) is uniform in  $t, s$ , which can be seen by a careful check throughout the whole proof process. We leave it to the reader.  $\square$

Combing the finite dimensional convergence and the tightness we finally complete the proof of Theorem 1.8.

## 7. APPENDIX

In this appendix, we present some existing results on the local behavior of the spectrum of Wigner matrices satisfying *Condition  $\mathbf{C}_0$* , which can be found in the recent work on the universality property of RMT. One can refer to the series [15]-[18] and [30]-[33] for instance. We also remark here the results stated below may be proved in their original articles under weaker conditions than those made in our paper. For ease of presentation, we reformulate them under the *Condition  $\mathbf{C}_0$*  without further explanation.

**Lemma 7.1.** *Let  $M_n$  be a Wigner matrix satisfying the Condition  $\mathbf{C}_0$ , one has for any interval  $I$  with its length  $|I| \geq n^{-1+Ac_0}$  for some sufficiently large constant  $A > 0$*

$$|N_n(I) - n \int_I \rho_{sc}(x) dx| \leq \varepsilon n |I| \quad (7.1)$$

with overwhelming probability. Moreover, one has

$$\sup_{1 \leq j \leq n} \|u_j(W_n)\|_{l^\infty} \leq n^{-1/2+\mathcal{O}(c_0)} \quad (7.2)$$

with overwhelming probability. Here  $u_j(W_n)$  is the unit eigenvector corresponding to  $\lambda_j(W_n)$ .

*Proof.* See Theorem 1.8 and Proposition 1.10 of [31] for instance.  $\square$

The second main lemma is an explicit description on the location of the eigenvalues proved in [18], named as the rigidity property for eigenvalues.

**Lemma 7.2** (Rigidity of eigenvalues). *Suppose that  $M_n$  is a Wigner matrix obeying the Condition  $\mathbf{C}_0$ . One has for some positive constants  $C, C', c$*

$$\begin{aligned} & P(\exists j : \lambda_j - \gamma_{j/n} \geq (\log n)^{C \log \log n} [\min(j, n-j+1)]^{-1/3} n^{-2/3}) \\ & \leq C' \exp[-(\log n)^{c \log \log n}] \end{aligned} \quad (7.3)$$

for  $n$  large enough. Moreover, one has

$$\mathbb{P}(\sup_{|x| \leq 5} n |F_n(x) - F_{sc}(x)| \geq (\log n)^{C \log \log n}) \leq C' \exp[-(\log n)^{c \log \log n}]$$

for sufficiently large  $n$ .

**Remark 7.3.** *It is not difficult to deduce (7.1) in Lemma 7.1 from (7.3) indeed. Above we state (7.1) separately just for convenience.*

The third lemma we need is the following large deviation estimate on the extreme eigenvalue of Wigner matrices under the assumption  $\mathbf{C}_0$ .

**Lemma 7.4.** *Suppose that  $M_n$  is a Wigner matrix obeying the Condition  $\mathbf{C}_0$ , one has*

$$\max_l |\lambda_l(M_n)| \leq 2 + \epsilon$$

*with overwhelming probability for any fixed  $\epsilon > 0$ . Moreover, there exists for any  $K \geq 3$*

$$\mathbb{P}(\max_l |\lambda_l(M_n)| \geq K) \leq \exp(-cn^c \log K)$$

*for sufficiently large  $n$*

*Proof.* Observe that the first statement is a direct consequence of Lemma 7.2. And the second one was proved by Erdős, Yau and Yin in [17] (See Lemma 7.2 of [17]).  $\square$

The last one is a uniform variance estimate for any single eigenvalue in the bulk, which was proved recently by Dallaporta in [9].

**Lemma 7.5** (eigenvalue variance bounds, [9]). *Let  $M_n$  be a complex Wigner matrices satisfying Condition  $\mathbf{C}_0$ , then for any  $0 < \delta \leq \frac{1}{2}$ , there exists a constant  $C(\delta) > 0$  such that for all  $n \geq 2$  and  $\delta n \leq l \leq (1 - \delta)n$ ,*

$$\text{Var}\{\lambda_l\} \leq C(\delta) \frac{\log n}{n^2}.$$

## REFERENCES

- [1] T.W. Anderson. Asymptotic theory for principal component analysis. Ann. Math. Statist. Vol. 34, No. 1, 122-148 (1963)
- [2] Z.D. Bai, J. Silverstein. Spectral analysis of large dimensional random matrices, Science Press. (2006)
- [3] Z.D. Bai, X.Y. Wang and W. Zhou. CLT for linear spectral statistics of Wigner matrices. Electron. J. Probab. 14, No.83, 2391-2417 (2009)
- [4] Z.D. Bai, J.F. Yao. On the convergence of the spectral empirical process of Wigner matrices. Bernoulli, 11, No.6, 1059-1092 (2005)
- [5] Z. Bai, M. Fahey, G. Golub, M. Menon, and E. Richter. Computing partial eigenvalue sum in electronic structure calculations. Technical Report, SCCM-98-03, Stanford University, Jan 1998.
- [6] P. Billingsley. Convergence of Probability Measures, Wiley, New York (1968)
- [7] S. Chatterjee. Fluctuations of eigenvalues and second order Poincaré inequalities. Probab. Theory Related Fields, Vol. 143, No. 1-2, 1-40 (2009)
- [8] O. Costin, J.L. Lebowitz. Gaussian fluctuations in random matrices. Phys. Rev. Lett. 75, p. 69-72 (1995)
- [9] S. Dallaporta. Eigenvalue variance bounds for Wigner random matrices. Preprint, arXiv: 1203.1597v2 (2012)
- [10] S. Dallaporta, V. Vu. A note on the central limit theorem for the eigenvalue counting function of Wigner matrices. Electron. Commun. Probab. 16, No. 30, 214-322 (2011)
- [11] E.B. Davies. Spectral Theory and Differential Operators. Cambridge Studies in Advanced Mathematics 42. Cambridge, UK: Cambridge University Press. (1995)

- [12] P. Deift. Orthogonal Polynomials and Random Matrices: A Riemann-Hilbert Approach. Courant Lecture Notes in Mathematics 18, American Mathematical Society, Providence, RI. (2009)
- [13] A. Edelman, T. Arias, and T. Smith. The geometry of algorithms with orthogonality constraints. SIAM J. on Mat. Anal. Appl. 20, No.2, 303-353 (1998)
- [14] N.M. Ercolani, K.D.T.-R. McLaughlin. Asymptotics of the partition function for random matrices via Riemann-Hilbert techniques and applications to graphical enumeration. IMRM, No. 14, 755-820(2003)
- [15] L. Erdős, B. Schlein, and H.-T. Yau. Semicircle law on short scales and delocalization of eigenvectors for Wigner random matrices. Ann. Probab. 37, 815-852 (2009)
- [16] L. Erdős, J.A. Ramírez, B. Schlein, and H.-T. Yau. Universality of sine-kernel for Wigner matrices with a small Gaussian perturbation. Electron. J. Probab. 15, No.18, 526-603 (2010).
- [17] L. Erdős, H.-T. Yau, and J. Yin. Bulk universality for generalized Wigner matrices. Proba. theory and related fields, in press, arXiv: 1001.3453 (2010)
- [18] L. Erdős, H.-T. Yau, and J. Yin. Rigidity of eigenvalues of generalized Wigner matrices. Advances in Mathematics. 229, No.3, 1435-1515 (2011)
- [19] F. Götze, A.N. Tikhomirov. The rate of convergence for spectra of GUE and LUE matrix ensembles. Cent. Eur. J. Math. 3, No.4, 666-704 (2005)
- [20] J. Gustavsson. Gaussian fluctuations of eigenvalues in the GUE. Ann. Inst. H. Poincar. Probab. Statist. 41, No.2, 151-178 (2005)
- [21] R.A. Horn, C.R. Johnson. Matrix Analysis, Cambridge University Press, New York. (1985)
- [22] A. Lytova, L. Pastur. Central limit theorem for linear eigenvalue statistics of random matrices with independent entries. Ann. of Probab. 37, No.5, 1778-1840. (2009)
- [23] M.L. Mehta. Random matrices. Second edition. Academic, Boston. (1991)
- [24] S. Oymak, B. Hassibi. New null space results and recovery thresholds for matrix rank minimization. Preprint, arXiv:1011.6326 (2010)
- [25] L. Pastur. Limiting laws of linear eigenvalue statistics for Hermitian matrix models. J. Math. Phys. 47 No. 10, 103303. (2006)
- [26] S. O'Rourke, D. Renfrew and A. Soshnikov. On Fluctuations of Matrix Entries of Regular Functions of Wigner Matrices with Non-identically Distributed Entries. J. Theor. Probab. DOI: 10.1007/s10959-011-0396-x (2011)
- [27] M. Shcherbina. Central limit theorem for linear eigenvalue statistics of the Wigner and sample covariance random matrices. Journal of Mathematical Physics, Analysis, Geometry, 7, No.2, 176-192 (2011)
- [28] Y. Sinai, A. Soshnikov. Central limit theorem for traces of large random symmetric matrices with independent matrix elements. Bol. Soc. brasil. Mat. (N.S.) 29, 1-24 (1998)
- [29] A. Soshnikov. Gaussian limit for determinantal random point fields. Ann. of Prob. vol. 30, No. 1, 171-187 (2002)
- [30] T. Tao, V. Vu. Random matrices: Universality of local eigenvalue statistics. Acta Math. 206, No. 1, 127-204 (2011)
- [31] T. Tao, V. Vu. Random matrices: Universality of local eigenvalue statistics up to the edge. Comm. Math. Phys. vol 298, No.2, 549-572 (2010)
- [32] T. Tao, V. Vu. A central limit theorem for the determinant of a Wigner matrix. arxiv: 1111.6300v3(2011)
- [33] T. Tao, V. Vu. Random matrices: Sharp concentration of eigenvalues. arXiv: 1201.4789v1 (2012)

DEPARTMENT OF MATHEMATICS, ZHEJIANG UNIVERSITY, P. R. CHINA  
*E-mail address:* zhigangbao@zju.edu.cn

DIVISION OF MATHEMATICAL SCIENCES, SCHOOL OF PHYSICAL AND MATHEMATICAL SCIENCES,  
 NANYANG TECHNOLOGICAL UNIVERSITY, SINGAPORE 637371  
*E-mail address:* gmpan@ntu.edu.sg

DEPARTMENT OF STATISTICS AND APPLIED PROBABILITY, NATIONAL UNIVERSITY OF SINGAPORE, SINGAPORE 117546

*E-mail address:* `stazw@nus.edu.sg`